Editorial

Occupational Therapy in the Age of Artificial Intelligence: Preserving Human Occupation While Embracing Digital Innovation

Allied Medical Research Journal
Vol:4 No.3: 01-02
© The Author (s) 2026

Neelum Zehra

Ziauddin College of Occupational Therapy, Ziauddin University, Karachi, Pakistan

Occupational therapy (OT) has long been recognized as a profession centered on enabling individuals to participate in meaningful occupations that promote health, well-being, and quality of life. Built upon the principles of person-centered care, occupational therapy emphasizes the therapeutic value of daily activities while considering the physical, psychological, social, and environmental factors influencing occupational performance. As healthcare undergoes rapid technological transformation, the profession now stands at a pivotal moment where artificial intelligence (AI), digital health technologies, and data-driven rehabilitation are reshaping clinical practice. Rather than viewing these innovations as a threat, occupational therapists have an opportunity to redefine their role in an increasingly digital healthcare ecosystem.

AI has already begun influencing rehabilitation through automated assessments, predictive analytics, virtual rehabilitation platforms, wearable sensors, and intelligent assistive technologies. Machine learning algorithms can analyze movement patterns, identify subtle functional impairments, and monitor patient progress with unprecedented precision. Smart homes equipped with AI-enabled devices can support older adults in maintaining independence, while virtual reality systems provide engaging therapeutic environments for individuals recovering from neurological and orthopedic conditions.

Despite these technological advances, the essence of occupational therapy remains irreplaceably human. Occupation extends beyond measurable physical performance; it encompasses identity, purpose, culture, values, motivation, and participation in meaningful life roles. These dimensions cannot be fully captured by algorithms alone. An individual may demonstrate adequate physical function through digital assessment yet continue to experience occupational deprivation due to social isolation, environmental barriers, or psychological challenges. Occupational therapists possess the unique expertise to interpret these complexities within the broader context of a person's life. The integration of AI should therefore be viewed as an enhancement rather than a

Email: neelum.zehra@zu.edu.pk

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

replacement of occupational therapy practice.

By automating routine documentation, outcome measurement, and data analysis, AI can reduce administrative burden and allow therapists to dedicate more time to therapeutic reasoning, collaborative goal setting, and individualized intervention planning. Technology should augment clinical judgment while preserving the therapeutic relationship that lies at the heart of occupational therapy.

Another emerging consideration is the ethical use of AI in rehabilitation. Issues related to data privacy, algorithmic bias, accessibility, and digital equity require careful attention. AI systems trained on limited populations may not accurately represent individuals from diverse cultural or socioeconomic backgrounds, potentially widening existing healthcare disparities. Occupational therapists, with their commitment to occupational justice and client-centered practice, are well positioned to advocate for equitable and inclusive implementation of digital technologies.

AI also opens new opportunities for research within occupational therapy. Large datasets collected through wearable devices and digital platforms can facilitate longitudinal monitoring of occupational performance, enabling researchers to identify patterns of recovery and personalize interventions. Predictive models may assist clinicians in identifying patients at risk of functional decline, supporting early intervention and preventive care. However, rigorous validation studies are necessary to ensure that AI-generated recommendations align with occupational therapy principles and improve meaningful participation rather than solely clinical metrics. Importantly, the profession must continue to emphasize that technology is a means to support occupation not an end in itself. Successful rehabilitation should ultimately be measured by an individual's ability to engage in meaningful life activities, participate within their communities, and achieve personal goals. AI can provide valuable information, but occupational therapists remain essential in translating that information into interventions that respect each person's unique occupational identity.

As healthcare moves toward precision rehabilitation and digitally integrated care, occupational therapy has an opportunity to lead rather than follow technological innovation. By embracing AI responsibly while preserving the profession's humanistic foundations, occupational therapists can ensure that technological progress enhances not diminishes the meaning of occupation. The future of occupational therapy lies not in choosing between technology and humanity, but in thoughtfully integrating both to create more accessible, personalized, and effective rehabilitation services.

The profession has consistently evolved in response to societal change. The emergence of AI represents another milestone in that evolution, challenging occupational therapists to innovate while remaining steadfast in their commitment to meaningful occupation, inclusion, and holistic care. If embraced with ethical responsibility and clinical wisdom, AI has the potential to strengthen occupational therapy's contribution to global health and redefine rehabilitation for the decades ahead.

Original Article

Allied Medical Research Journal
Vol:4 No.3: 03-12
© The Author (s) 2026

Comparative Effects of Mirror Therapy, Proprioceptive Neuromuscular Facilitation, and Constraint-Induced Movement Therapy on Upper Limb Motor Recovery in Stroke Survivors: A Three-Arm Randomized Controlled Trial

Subhan Ali Gondal[1], Aliha Shafqat[2], Namara Atif[2], Saleha Khan[2], Aqsa Noreen[3], Nusrat Naseem[4]

[1]*Ikram Hospital, Gujrat, Pakistan*
[2]*Physio Care Centre, Gujrat, Pakistan*
[3]*Mukabbir University of Science & Technology, Gujrat, Pakistan*
[4]*Health Services Academy, Islamabad, Pakistan*

ABSTRACT

Background: The majority of stroke survivors have upper-limb weakness, which reduces their functional independence in the real world, but there is limited evidence from trials that directly compare upper-limb rehabilitation interventions. The study compared the efficacy of three types of therapies (Mirror Therapy (MT), Proprioceptive Neuromuscular Facilitation (PNF), and Constraint-Induced Movement Therapy (CIMT)) in restoring motor function of the affected upper limb after stroke, when combined with conventional physiotherapy.

Methods: This was a three-arm parallel-group trial of 90 adults with post-stroke upper-limb hemiparesis, with blinded outcome assessment. The participants were allocated in a 1:1:1 ratio to MT, PNF, or CIMT, each with conventional physiotherapy, 6 days per week. The Action Research Arm Test (ARAT), Fugl-Meyer Assessment of the Upper Extremity (FMA-UE), Modified Ashworth Scale (MAS), and grip strength were measured by a blinded assessor at baseline and at 12 weeks.

Results: All 90 participants completed the trial with significant improvements from baseline on all four measures ($p<0.001$). A between-group difference at week 12 emerged only for the ARAT (one-way ANOVA, $F=9.335$, $p<0.001$): CIMT scored higher than both MT ($p=0.002$) and PNF ($p<0.001$), whereas MT and PNF were statistically indistinguishable. FMA-UE, MAS, and grip strength showed no between-group differences.

Conclusion: All three interventions improved upper-limb function with CIMT showing the greatest benefit in task performance, as measured by the ARAT. When CIMT is not feasible, MT and PNF are effective alternatives. This highlights the need to tailor therapy choices to each patient's goals and resources.

Keywords: *Motor Recover, Quality of Life; Rehabilitation; Stroke; Upper Extremity.*

Received: April 29th, 2026; **Revised:** June 25th, 2026; **Accepted:** July 9th, 2026

Corresponding Email: dr.subhangondal@gmail.com

DOI: https://doi.org/10.59564/amrj/04.03/001

INTRODUCTION

Worldwide, stroke ranks among the foremost causes of long-term disability, with over 12 million new cases each year and more than 100 million survivors living with lasting impairment[1]. Global burden-of-disease estimates likewise rank stroke among the leading contributors to disability-adjusted life years across all age groups[2]. The upper limb is affected more often than any other body region after stroke, with weakness present in roughly 55-85% of survivors during both the acute and chronic phases[3-4]. This typically appears as weakness, altered muscle tone, impaired coordination, and reduced dexterity, all of which limit a person's capacity to manage daily activities independently and erode quality of life[3]. Despite improved survival after acute stroke, only a small proportion of patients regain full use of the paretic arm, underscoring the ongoing need for effective, evidence-based approaches to upper limb rehabilitation[5].

A range of neurorehabilitation approaches, each grounded in a distinct neurophysiological rationale, is used to promote upper-limb recovery after stroke. Mirror therapy (MT) provides visual feedback that creates the illusion of normal movement in the affected

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

limb, a mechanism thought to drive cortical reorganization and motor relearning[6]. Trials of MT, whether delivered conventionally or through virtual-reality enhancement, have repeatedly shown clinically meaningful gains in motor function, sensorimotor integration, and daily functioning when added to conventional therapy[6-9]. PNF instead relies on diagonal and spiral movement patterns, paired with proprioceptive stimulation, to enhance neuromuscular efficiency and reduce spasticity, and has been shown to improve motor control and functional independence in stroke survivors[10]. Among the most extensively studied approaches is constraint-induced movement therapy (CIMT), which restrains the unaffected limb to force intensive, repetitive practice with the paretic limb, thereby reversing learned non-use through use-dependent cortical plasticity[11-13]. Each of these three approaches outperforms conventional physiotherapy alone, though they act on distinct facets of neuroplasticity, motor learning, and behavioral adaptation.

Although evidence supports the efficacy of each approach individually, few trials have directly compared MT, PNF, and CIMT head-to-head within a single, adequately powered study. Most existing research has evaluated these therapies in isolation or against conventional therapy alone rather than against one another, leaving clinicians without a solid evidence base for choosing among them for a given patient[14,15]. This gap matters most in resource-limited settings, where selecting the most time- and cost-efficient option can shape both patient recovery and resource use. Adding to the uncertainty, many earlier comparative studies have relied on small sample sizes, a single outcome measure, or short follow-up periods, limiting the extent to which their conclusions about long-term upper limb recovery can be generalized.

Closing this gap calls for a well-designed, adequately powered randomized controlled trial that compares these three widely used interventions head-to-head using standardized, validated outcome measures. This study was therefore undertaken to determine how mirror therapy, proprioceptive neuromuscular facilitation, and constraint-induced movement therapy, each added to conventional physiotherapy, affect upper-limb motor recovery, functional performance, spasticity, and grip strength in people with post-stroke upper-limb hemiparesis.

METHODOLOGY

Study Design and Setting

This study was designed as a three-arm, parallel-group, assessor-blinded randomized controlled trial conducted in the Department of Physiotherapy, Ikram Hospital, Gujarat, Punjab, Pakistan, between July 2025 and January 2026. Eligible participants with post-stroke upper-limb hemiparesis were randomly allocated in equal numbers to one of three groups; Mirror Therapy (MT), Proprioceptive Neuromuscular Facilitation (PNF), or Constraint-Induced Movement Therapy (CIMT); each delivered alongside conventional physiotherapy.

Calculation of Sample Size

The number of samples was determined based on a previous trial comparing CIMT versus MT for hand function recovery after stroke, which found between-group Fugl-Meyer Assessment (FMA) scores of 27.461 ± 2.757 for the CIMT group and 24.923 ± 2.253 for the MT group[16]. The number of subjects needed was estimated using the standard formula for comparing two independent means:

$$n = 2\sigma^2(Z_{\alpha/2} + Z_{\beta})^2 / (\mu1 - \mu2)^2$$

Where:

- $Z_{\alpha/2} = 1.96$ (95% confidence level, two-tailed)
- $Z_{\beta} = 0.84$ (80% statistical power)
- σ = pooled standard deviation of the two groups
- $(\mu1 - \mu2)$ = difference between group means

This indicated that each group should have about 16 participants, based on a pooled SD of 2.518 and a mean difference between groups of 2.538. This number was increased to 30 participants per group to account for drop-off and increase statistical power, for a total of 90 participants across the three arms.

Randomization and Allocation Concealment

A computer-generated random number sequence was used to assign eligible participants to one of the three intervention groups (MT, PNF, or CIMT). To conceal allocation, sequentially numbered, opaque, sealed envelopes (SNOSE) were prepared in advance and opened only after a participant had enrolled and completed baseline assessment; this step was carried out by an investigator with no role in delivering treatment or assessing outcomes.

Participants

Ninety participants with post-stroke upper-limb hemiparesis, aged 25–65 years, of either sex, were recruited from the Department of Physiotherapy, Ikram Hospital, Gujrat, Pakistan.

Inclusion Criteria:

- Clinically or radiologically confirmed stroke patients
- Both male and female participants
- Age range of 25–65 years
- Presence of unilateral upper limb hemiparesis

Exclusion Criteria:

- Inability to perform prescribed exercises
- Lack of willingness or denial to provide informed consent
- Cognitive impairment has already been diagnosed, affecting comprehension or protocol compliance

Interventions

All three groups received their assigned intervention in addition to conventional physiotherapy (passive/active range-of-motion exercises, positioning, and functional mobility training), delivered five sessions per week for six weeks, with outcome assessment repeated at the 12-week mark.

- **Mirror Therapy (MT) Group:** Participants were seated in front of a mirror box oriented in the mid-sagittal plane with the unaffected limb positioned in front of the mirror facing the participant and the affected limb behind the mirror. The participants performed a series of movements with the non-affected hand while observing the reflection of their own hand; this mechanism produced the illusion of normal motor function in the affected hand. Sessions lasted approximately 30 minutes and included 10-15 repetitions of each movement pattern, with rest breaks to reduce fatigue. It has been hypothesized that this visual input activates mirror-neuron networks and activates the premotor cortex in the affected side, thus promoting motor relearning and reorganization[6].

- **Proprioceptive Neuromuscular Facilitation (PNF) Group:** Participants were given manually resisted exercises based on Kabat's diagonal movement patterns (D1 and D2 Flexion/Extension patterns of the upper limb) as well as facilitation techniques such as rhythmic initiation, slow reversal, hold-relax, and contract-relax. The number of repetitions begun at each diagonal pattern was 10-15 per set, for a total of approximately 30 minutes of each session, and manual proprioceptive input (stretch, traction, and approximation) was applied to the shoulder, elbow, and wrist to maximize the recruitment of neuromuscular units, improve coordination, and decrease spasticity via reciprocal inhibition, respectively[10].

- **Constraint-Induced Movement Therapy (CIMT) Group:** Therapists prescribed wearing a padded mitt to restrict use of the non-affected upper limb for approximately 90% of waking hours during treatment and, if possible, during structured CIMT at home. Supervised sessions lasted approximately 30 minutes. They included intensive, repetitive, task-oriented practice with the affected limb using the "shaping" technique, which involves dividing functional tasks (e.g., reaching, grasping, transferring objects, and manipulating daily items) into successive approximations of increasing difficulty. The forced and repetitive use of the paretic limb is designed to reverse learned non-use and foster use-dependent reorganization of the cortex[11,12].

Outcome Measures

All outcome measures were recorded at baseline and after 12 weeks of intervention by a blinded assessor not involved in treatment delivery.

- **Action Research Arm Test (ARAT):** This is an observational test that consists of 19 items across four subtests (pinch, grasp, gross movement, and grip) with each subtest scored on an ordinal scale from 0 (no movement at all) to 3 (normal movement), for a maximum of 57 points, with higher scores indicating a better upper arm function. Both the Fugl-Meyer scale and the ARAT have good criterion validity in post-stroke patients with known motor impairments, and the ARAT exhibits excellent internal consistency across its items in this patient population[17].

- **Fugl-Meyer Assessment for Upper Extremity (FMA-UE):** This is a 33 Item performance-based assessment of motor impairment involving reflexes, movement within and outside of synergy, wrist and hand function, and movement speed and coordination. They are rated on a 3-point

scale (0 = unable to perform, 1 = partial performance, 2 = full performance), with a maximum possible score of 66, where a higher score indicates less impairment. This is considered the gold-standard stroke rehabilitation research tool because of its reliability, validity, and responsiveness[18].

- **Modified Ashworth Scale (MAS):** This is a clinician-rated spasticity test designed to assess resistance to passive stretch of the shoulder adductors, elbow flexors, and wrist flexors. A 6-point scale (0, no increase in tone; 1, 1+, 2, 3, 4, affected part rigid in flexion or extension) is used; the modal grade is recorded from three passive movement trials (one per muscle group), and higher scores indicate greater spasticity. This is because its standardized scoring protocol has demonstrated good-to-excellent reliability, both within and between the scorers[19].

- **Grip Strength:** The maximal isometric grip strength of the affected hand was measured with the hand-held dynamometer setup when the participant sat with the shoulder adducted, elbow flexed at 90°, and forearm neutral. The final value was the mean of three trials with a short break between trials. This has been shown to have high test–retest and inter-rater reliability (ICC > 0.9) among stroke patients[20].

Statistical Analysis

SPSS (version 24.0) was used to analyze the data. Descriptive statistics (mean ± SD) were calculated for each outcome variable, and the Shapiro-Wilk test was performed to assess normality. The one-way analysis of variance (ANOVA) was used with Tukey's post hoc test for pairwise comparisons to compare the groups at baseline and at the end of the intervention. Changes within groups over time (baseline to week 12) were evaluated using paired t-tests or the Wilcoxon signed-rank test, depending on the data distribution. All outcome tabulations were presented as 95% confidence intervals (CIs), and statistical significance was defined as $p<0.05$.

Ethical Considerations

Approval for this study was obtained from the Institutional Review Board of Ikram Hospital, Gujrat (IRB No. IKG-002/24) prior to recruitment and registered prospectively with ClinicalTrials.gov (Registration No. NCT07683130).All procedures adhered to the ethical principles set out in the Declaration of Helsinki and its subsequent amendments. Each participant, or their legal attendant where relevant, provided written informed consent after receiving a complete explanation of the study procedures and their potential risks and benefits, and all participants were informed that they could withdraw at any time without affecting their ongoing clinical care.

RESULTS

Participant Flow and Baseline Characteristics

A total of 90 participants with post-stroke upper-limb hemiparesis were enrolled and evenly distributed across the Mirror Therapy (MT, n=30), Proprioceptive Neuromuscular Facilitation (PNF, n=30), and Constraint-Induced Movement Therapy (CIMT, n=30) groups, each receiving their assigned intervention alongside conventional physiotherapy. Every participant completed both the six-week intervention and the 12-week follow-up assessment, and none dropped out of any arm. The three groups were similar at baseline in demographic and clinical characteristics (Table 1); one-way ANOVA found no significant differences in age ($F = 0.20$, $p = 0.821$) or time since stroke ($F = 1.18$, $p = 0.311$), supporting adequate randomization, and gender, affected side, and stroke type were likewise distributed similarly across groups.

Table 1: Baseline demographic and clinical characteristics of the three treatment groups.

Characteristic	Mirror Therapy (n=30)	PNF (n=30)	CIMT (n=30)	p-value
Age, years (mean ± SD)	46.20 ± 8.12	46.83 ± 9.77	47.83 ± 8.52	0.821†
Time since stroke, months (mean ± SD)	7.22 ± 3.10	6.63 ± 3.12	6.11 ± 3.37	0.311†
Gender, Male / Female, n	16 / 14	18 / 12	18 / 12	
Affected side, Left / Right, n	15 / 15	17 / 13	15 / 15	
Stroke type, Ischemic / Hemorrhagic, n	19 / 11	28 / 2	23 / 7	

†One-way ANOVA

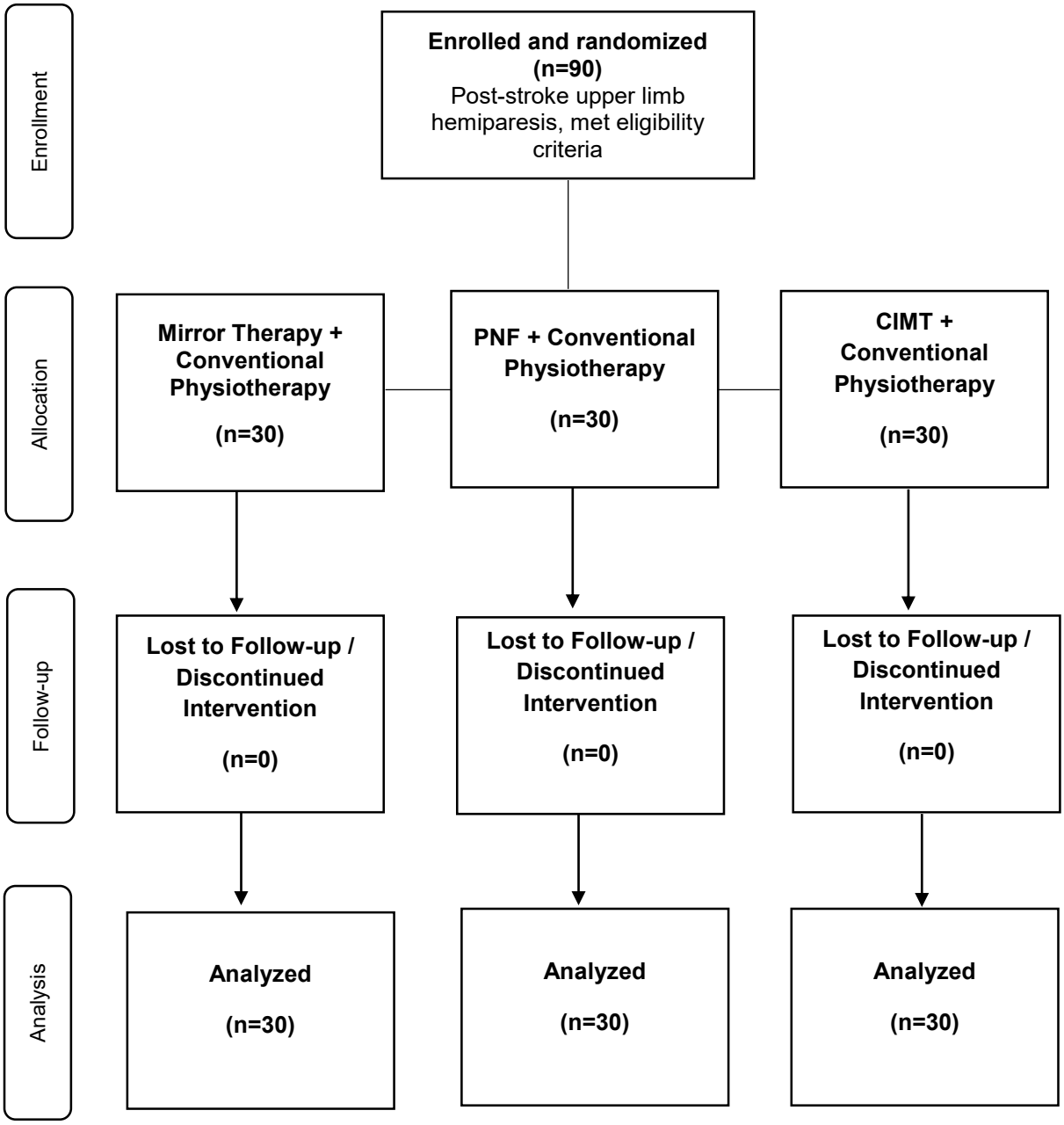

Figure 1. CONSORT diagram showing participant enrollment, group allocation, follow-up, and analysis

Baseline Comparability of Outcome Measures

None of the four outcome measures differed significantly between groups at baseline: ARAT (p=0.175), FMA-UE (p=0.487), MAS (p=0.483), and grip strength (p=0.689) (Table 2), one-way ANOVA, confirming that the three groups were closely matched on motor impairment and function before treatment began.

Table 2: Baseline and week-12 scores, within-group change, and between-group comparison for all outcome measures, by treatment group.

Outcome measure	Group	Baseline (mean ± SD)	Week 12 (mean ± SD)	Mean change ± SD	Within-group p-value	Between-group, baseline	Between-group, week 12
ARAT (0–57)	Mirror Therapy	20.73 ± 7.41	29.17 ± 6.81	8.43 ± 3.19	<0.001	F = 1.777, p = 0.175	F = 9.335, p < 0.001
	PNF	21.37 ± 7.63	28.47 ± 8.67	7.10 ± 3.33	<0.001		
	CIMT	23.93 ± 5.69	36.10 ± 7.08	12.17 ± 4.84	<0.001		

FMA-UE (0–66)	Mirror Therapy	30.97 ± 9.56	43.20 ± 9.05	12.23 ± 3.46	<0.001	F = 0.725, p = 0.487	F = 0.105, p = 0.900
	PNF	32.97 ± 10.60	42.60 ± 11.39	9.63 ± 5.04	<0.001		
	CIMT	30.03 ± 8.66	43.87 ± 11.45	13.83 ± 5.64	<0.001		
MAS (0–4)‡	Mirror Therapy	1.20 ± 0.90	0.98 ± 0.88	−0.22 ± 0.39	0.005	F = 0.735, p = 0.483	F = 0.788, p = 0.458
	PNF	1.28 ± 0.99	1.13 ± 0.96	−0.15 ± 0.33	0.017		
	CIMT	1.47 ± 0.71	1.27 ± 0.78	−0.20 ± 0.39	0.008		
Grip strength (kg)	Mirror Therapy	9.04 ± 3.40	11.89 ± 3.62	2.85 ± 1.12	<0.001	F = 0.374, p = 0.689	F = 1.337, p = 0.268
	PNF	8.33 ± 3.63	10.97 ± 3.85	2.65 ± 1.01	<0.001		
	CIMT	8.49 ± 2.91	12.42 ± 2.82	3.92 ± 1.17	<0.001		

‡MAS grades were coded numerically (0, 1, 1+ = 1.5, 2, 3, 4) for statistical analysis. Within-group p-values are from paired t-tests (ARAT, FMA-UE, grip strength) or the Wilcoxon signed-rank test (MAS); between-group p-values are from one-way ANOVA (df = 2,87).

‡MAS grades were coded numerically (0, 1, 1+ = 1.5, 2, 3, 4) for statistical analysis.

Action Research Arm Test (ARAT)

From baseline to week 12, all three groups showed significant within-group gains in ARAT score (paired t-test, $p < 0.001$ for each group; Table 2, Figure 2). The CIMT group gained the most (12.17 ± 4.84 points), ahead of Mirror Therapy (8.43 ± 3.19 points) and PNF (7.10 ± 3.33 points). One-way ANOVA identified a significant between-group difference in ARAT score at week 12 ($F = 9.335$, $p < 0.001$).

Post hoc Tukey testing confirmed that CIMT scored significantly higher on the ARAT than Mirror Therapy (mean difference = 6.93, 95% CI 2.28–11.59, $p = 0.002$) and PNF (mean difference = 7.63, 95% CI 2.98–12.29, $p < 0.001$), whereas Mirror Therapy and PNF did not differ significantly from one another (mean difference = 0.70, 95% CI −3.96–5.36, $p = 0.932$) (Table 3, Figure 2).

Table 3: Tukey's post hoc pairwise comparison for ARAT at week 12, the only outcome measure showing a significant overall between-group effect.

Group-wise comparison (ARAT, Week 12)	Mean difference	95% CI	p-adj (Tukey)
CIMT vs Mirror Therapy	6.93	2.28 to 11.59	0.002**
CIMT vs PNF	7.63	2.98 to 12.29	<0.001***
Mirror Therapy vs PNF	0.70	−3.96 to 5.36	0.932

***p < 0.01, ***p < 0.001.*

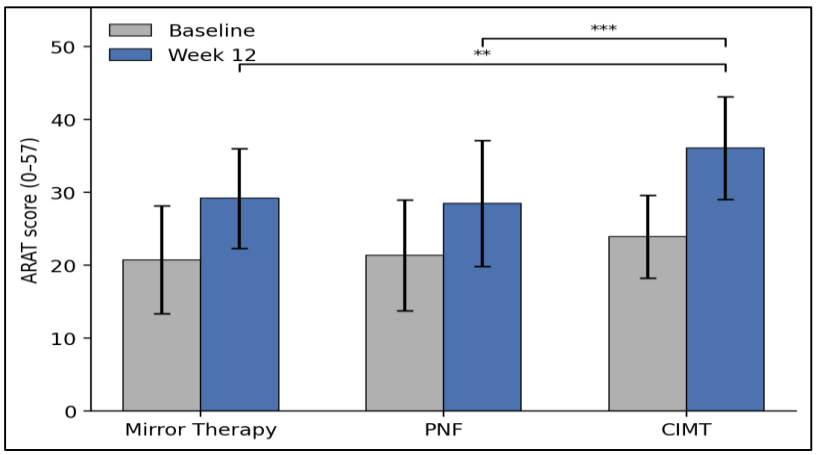

Figure 2: Mean ARAT scores (± SD) at baseline and week 12 across the Mirror Therapy, PNF, and CIMT groups. **p < 0.01, ***p < 0.001 (Tukey post hoc, Week-12 comparison vs CIMT).

Fugl-Meyer Assessment – Upper Extremity (FMA-UE)

All three groups improved significantly on the FMA-UE between baseline and week 12 (paired t-test, $p < 0.001$ for each group; Table 2, Figure 3), with average gains of 13.83 ± 5.64 points for CIMT, 12.23 ± 3.46 points for Mirror Therapy, and 9.63 ± 5.04 points for PNF. The between-group comparison at week 12 was not significant ($F = 0.105$, $p = 0.900$), meaning the three interventions led to similar reductions in motor impairment as captured by the FMA-UE.

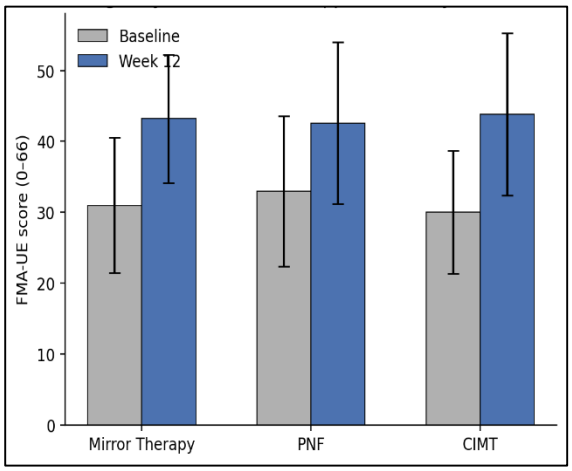

Figure 3: Mean FMA-UE scores (± SD) at baseline and week 12 across the three groups. One-way ANOVA showed no difference in the results between the groups at week 12 ($p = 0.900$).

Modified Ashworth Scale (MAS)

Spasticity, measured by the MAS, decreased significantly from baseline to week 12 in all three groups (Wilcoxon signed-rank test; $p = 0.008$ for CIMT, $p = 0.005$ for Mirror Therapy, $p = 0.017$ for PNF; Table 2), though the reduction was modest in every group. No between-group difference emerged at week 12 ($F = 0.788$, $p = 0.458$), pointing to a similar degree of spasticity reduction across all three treatment approaches.

Grip Strength

Maximal isometric grip strength of the affected hand rose significantly over the 12 weeks in every group (Table 2, paired t-test, $p < 0.001$ in each group). CIMT produced the largest within-group gain (3.92 ± 1.17 kg), ahead of Mirror Therapy (2.85 ± 1.12 kg) and PNF (2.65 ± 1.01 kg). Although this pattern favored CIMT, the difference between groups at week 12 was not statistically significant ($F = 1.337$, $p = 0.268$).

DISCUSSION

This three-arm trial aimed to compare Mirror Therapy, Proprioceptive Neuromuscular Facilitation, and Constraint-Induced Movement Therapy for upper-limb motor recovery in people with post-stroke hemiparesis. Every intervention led to significant within-group gains across all four outcomes, ARAT, FMA-UE, MAS, and grip strength demonstrating that each approach can drive meaningful recovery over a 12-week course. However, only the ARAT, which captures performance on functional tasks, revealed a statistically significant difference between groups: CIMT improved more than either Mirror Therapy or PNF, while the latter two produced similar gains. FMA-UE, MAS, and grip strength showed no between-group differences, suggesting the three interventions had broadly comparable effects on motor impairment, spasticity, and strength.

CIMT's greater gains on the ARAT align with the theoretical rationale of forced-use therapy, which counters learned non-use by restraining the unaffected limb and requiring intensive, repetitive, task-oriented use of the paretic limb[11-13]. Since the ARAT closely tracks real-world manipulative actions such as grasping, gripping, and pinching, it may be especially responsive to the task-specific, functionally grounded practice built into CIMT's shaping technique, where movements are drawn from everyday activities rather than isolated exercises. This mirrors a recent three-arm RCT that, for the first time, directly compared CIMT and MT and likewise found a greater, though not exclusive, advantage for CIMT in functional strength and independence measures[21].

It also aligns with a network meta-analysis of upper limb exercise interventions after stroke, which found task-specific and constraint-based training to be the most effective for restoring upper limb motor function[22], and with a separate network meta-analysis of non-conventional upper limb therapies, in which CIMT-based approaches consistently ranked among the top performers for activity-level outcomes[23].

Unlike the ARAT results, the lack of a significant between-group difference on the FMA-UE suggests that MT, PNF, and CIMT converge on a common pathway to recovery in motor impairment, rather than diverge. Because the FMA-UE assesses reflex activity, synergy patterns, and isolated joint movement rather than integrated functional performance[18], it may respond more to the overall amount of active practice matched here at five sessions weekly for six weeks than to the particular training format (mirror feedback, proprioceptive facilitation, or forced use).

A comparative trial of modified CIMT, proprioceptive training, and task-oriented training reached a similar conclusion, reporting significant within-group gains across all three approaches with broadly equivalent outcomes, and suggesting that the amount and consistency of practice, rather than the specific technique used, drives recovery at the impairment level[24]. In line with this, a systematic review of dosing and timing in upper limb interventions after stroke found treatment intensity and duration to be the key drivers of motor recovery, regardless of which intervention was used[25].

Outcomes for Mirror Therapy and PNF were statistically indistinguishable across all four measures, echoing earlier comparative work indicating that these two therapies confer similar overall benefit despite working through different mechanisms[14,15]. MT is thought to aid recovery by engaging the mirror neuron system, recruiting ipsilateral motor pathways, and restoring interhemispheric connectivity between the two primary motor cortices, as a recent resting-state fMRI study has shown[26]. PNF, on the other hand, depends on manually resisted diagonal movement patterns and proprioceptive facilitation to boost neuromuscular recruitment and dampen spasticity via reciprocal inhibition[10]. The fact that these mechanistically distinct approaches yielded statistically equivalent clinical outcomes here implies that, at least in this population, several different neurorehabilitative routes can converge on similar functional gains when matched for intensity and duration.

Spasticity, as measured by the MAS, decreased significantly across all three groups, with no differences between them, indicating that adding MT, PNF, or CIMT to conventional physiotherapy produced a similarly modest anti-spastic effect. This aligns with a recent systematic review and meta-analysis reporting that MT reduces spasticity to a degree comparable to active exercise-based or traditional interventions, without clear evidence of superiority[27]. The modest average change in MAS grade seen across all three groups here may also reflect a ceiling effect, since baseline spasticity was already mild to moderate (1.20–1.47), leaving limited room for further reduction over the 12-week trial period.

Grip strength rose significantly across all groups, most markedly in the CIMT group, though the between-group difference fell short of significance. This is consistent with a modest, though not statistically significant, advantage for CIMT on hand-strength outcomes relative to MT, and aligns with evidence that hand-held dynamometry, while more reliable in stroke patients than in healthy controls, may lack the sensitivity to detect between-group differences of this magnitude without a larger sample or longer follow-up.

These findings align with earlier work reporting numerically greater hand-function gains with CIMT than with MT in cerebrovascular patients, without diminishing MT's value as a widely accessible, low-cost alternative[16]. Taken together with the present trial, the broader picture is that while CIMT can offer functional benefits, MT and PNF remain valuable options, particularly for patients unable to tolerate CIMT's demanding training schedule or restraint protocol. From a clinical and resource-allocation perspective, these findings carry implications for rehabilitation planning, particularly in resource-constrained settings such as the one in this study. Where the main rehabilitation goal is functional independence in activities of daily living (ADLs), CIMT may be the preferred choice; MT and PNF, being less equipment-, supervision-, and tolerability-intensive, may be better suited when CIMT is impractical due to severe impairment, poor adherence to restraint, or limited therapist availability[9,24].

Given that the three therapies produced equivalent effects on impairment, spasticity, and strength, therapy selection should hinge on patient preference, cognitive and physical tolerance, and available resources, rather than on the assumption that one approach is universally superior.

Strengths and Limitations

A three-arm randomized design, a sufficient sample size, standardized and validated outcome measures, blinded assessment, and full follow-up at 12 weeks without attrition further increase confidence in these results. However, there are some constraints to consider. First, the single-center design and the relatively homogeneous sociodemographic profile of participants may limit the generalizability of these results to other settings and populations. Second, because only outcome assessors were blinded, there is a possibility of performance bias if the therapist and/or participants know which group they were assigned to. Third, the 12-week endpoint reflects only short-term recovery, and longer follow-up would be required to determine whether the benefit of CIMT is maintained, diminishes, or equals that of MT and PNF over

time, as seen in similar trials conducted in the upper limb. Fourth, because no neurophysiological or neuroimaging studies were performed, differences between groups cannot be definitively attributed to distinct underlying mechanisms and may be due to differences in training specificity alone.

Further trials using functional MRI or other imaging methods, as in mechanism studies of MT, may help determine whether the functional benefit of CIMT is due to distinct patterns of brain reorganization or to targeted training. Lastly, the study did not measure cost-effectiveness or patient-reported quality of life, both of which are important when considering implementation in a real-world, resource-limited rehabilitation setting. Further studies should include multi-center designs with longer follow-up periods, economic analyses, and hybrid or sequential intervention protocols (e.g., CIMT after MT) to better balance the functional advantages of CIMT with the availability and tolerability of MT and PNF.

CONCLUSION

All three interventions (Mirror Therapy, CIMT, and Proprioceptive Neuromuscular Facilitation), plus standard physiotherapy, were shown to produce statistically significant improvements in upper-limb motor function, spasticity, and grip strength in patients with post-stroke upper-limb hemiparesis over 12 weeks. CIMT produced a significantly greater benefit than either Mirror Therapy or PNF in activity level function (ARAT) and similar benefits in motor impairment (FMA-UE), spasticity (MAS), and grip strength. Based on these results, CIMT is more effective in improving functional task performance; however, Mirror Therapy and PNF are effective and well-tolerated when the CIMT's restraint-based protocol is not suitable. Therefore, the choice of adjunct therapy should be based on the individual patient's clinical situation, functional goals, and available clinical resources, and should not be made on the assumption that any one technique is superior.

Author Contributions

SAG: Conception and Design, Supervision, and Manuscript Writing.
AS: Study Design, Data Collection, and Data Analysis.
NA: Literature Review, Data Interpretation, and Manuscript Preparation.
SK: Data Collection, Manuscript Editing, and Critical Review.
AN: Manuscript Writing and Final Approval.
NN: Supervision, Critical Review, and Final Approval.

Funding Source
No sources.

Conflict of Interest
No conflict of interests.

Data Availability
Data will be available on request.

Ethical Approval
This study was approved by the Institutional Review Board of Ikram Hospital, Gujrat (IRB No. IKG-002/24) and registered with ClinicalTrials.gov (Registration No. NCT07683130).

REFERENCES

1. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, Fisher M, Pandian J, Lindsay P. World Stroke Organization (WSO): global stroke fact sheet 2022. International journal of stroke. 2022 Jan;17(1):18-29. https://doi.org/10.1177/17474930211065917.
2. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, Abbasi-Kangevari M, Abbastabar H, Abd-Allah F, Abdelalim A, Abdollahi M. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The lancet. 2020 Oct 17;396(10258):1204-22. https://doi.org/10.1016/S0140-6736(20)30925-9.
3. Tang Q, Yang X, Sun M, He M, Sa R, Zhang K, Zhu B, Li T. Research trends and hotspots of post-stroke upper limb dysfunction: a bibliometric and visualization analysis. Frontiers in Neurology. 2024 Oct 2;15:1449729. https://doi.org/10.3389/fneur.2024.1449729.
4. Dalton EJ, Jamwal R, Augoustakis L, Hill E, Johns H, Thijs V, Hayward KS. Prevalence of arm weakness, pre-stroke outcomes and other post-stroke impairments using routinely collected clinical data on an acute stroke unit. Neurorehabilitation and Neural Repair. 2024 Feb;38(2):148-60. https://doi.org/10.1177/15459683241229676.
5. O'Flaherty D, Ali K. Recommendations for upper limb motor recovery: an overview of the UK and European rehabilitation after stroke guidelines (2023). InHealthcare 2024 Jul 18 (Vol. 12, No. 14, p. 1433). MDPI. https://doi.org/10.3390/healthcare12141433.
6. Wen X, Li L, Li X, Zha H, Liu Z, Peng Y, Liu X, Liu H, Yang Q, Wang J. Therapeutic role of additional mirror therapy on the recovery of upper extremity motor function after stroke: a single-blind, randomized controlled trial. Neural plasticity. 2022;2022(1):8966920. https://doi.org/10.1155/2022/8966920.
7. Okamura R, Nakashima A, Moriuchi T, Fujiwara K, Ohno K, Higashi T, Tomori K. Effects of a virtual reality-based mirror therapy system on upper extremity rehabilitation after stroke: a systematic review and meta-analysis of randomized controlled trials. Frontiers in neurology. 2024 Jan 8;14:1298291. https://doi.org/10.3389/fneur.2023.1298291.
8. Kim H, Kim J, Jo S, Lee K, Kim J, Song C. Video augmented mirror therapy for upper extremity rehabilitation after stroke: a randomized controlled trial. Journal of Neurology. 2023 Feb;270(2):831-42. https://doi.org/10.1007/s00415-022-11410-6.
9. Saragih ID, Priyanti RP, Batubara SO, Lee BO. Effects of mirror therapy on upper limb motor function of patients with stroke: A systematic review and meta-analysis of randomized controlled trials. Clinical

Rehabilitation. 2025 Jan;39(1):23-34. https://doi.org/10.1177/02692155241299211.
10. Alashram AR, Alghwiri AA, Padua E, Annino G. Efficacy of proprioceptive neuromuscular facilitation on spasticity in patients with stroke: A systematic review. Physical Therapy Reviews. 2021 May 4;26(3):168-76. https://doi.org/10.1080/10833196.2021.1892281.
11. Uswatte G, Taub E, Lum P, Brennan D, Barman J, Bowman MH, Taylor A, McKay S, Sloman SB, Morris DM, Mark VW. Tele-rehabilitation of upper-extremity hemiparesis after stroke: Proof-of-concept randomized controlled trial of in-home Constraint-Induced Movement therapy. Restorative Neurology and Neuroscience. 2021 Sep 20;39(4):303-18. https://doi.org/10.3233/RNN-201100.
12. Hu J, Liu PL, Hua Y, Gao BY, Wang YY, Bai YL, Chen C. Constraint-induced movement therapy enhances AMPA receptor-dependent synaptic plasticity in the ipsilateral hemisphere following ischemic stroke. Neural Regeneration Research. 2020 Aug 24;16(2):319. https://doi.org/10.4103/1673-5374.290900.
13. Ceylan CM, Şen Eİ, Karaagac T, Şahbaz T, Yalıman A. Effect of modified constraint-induced movement therapy on upper extremity function for stroke patients with right/left arm paresis: a single-blind randomized controlled trial. Ahi Evran Medical Journal. 2023 May 1;7(2):155-64. DOI: https://doi.org/10.46332/aemj.1120884.
14. Zhang Y, Xing Y, Li C, Hua Y, Hu J, Wang Y, Ya R, Meng Q, Bai Y. Mirror therapy for unilateral neglect after stroke: A systematic review. European Journal of Neurology. 2022 Jan;29(1):358-71. https://doi.org/10.1111/ene.15122.
15. Hsu HY, Kuo LC, Lin YC, Su FC, Yang TH, Lin CW. Effects of a virtual reality–based mirror therapy program on improving sensorimotor function of hands in chronic stroke patients: a randomized controlled trial. Neurorehabilitation and neural repair. 2022 Jun;36(6):335-45. https://doi.org/10.1177/15459683221081430.
16. Hooria S, Gondal MJ, Kabir A, Mazahir S, Khan A, Kamray R, Mehak S, Imtiaz R. Comparison of Constraint-Induced Movement Therapy vs Mirror Therapy Effect in Infarcted Cerebrovascular Patients for Improving Hand Function: CIMT vs Mirror Therapy for Hand Function in CVA. Journal of Health and Rehabilitation Research. 2024 Sep 6;4(3):1-5. DOI: https://doi.org/10.61919/jhrr.v4i3.1319.
17. Fernández-Solana J, Pardo-Hernández R, González-Bernal JJ, Sánchez-González E, González-Santos J, Soto-Cámara R, Santamaría-Pelaez M. Psychometric properties of the Action Research Arm Test (ARAT) scale in post-stroke patients—Spanish population. International journal of environmental research and public health. 2022 Nov 13;19(22):14918. https://doi.org/10.3390/ijerph192214918.
18. Hervé-Colas J, Newton SP, Engelter ST, Hayward KS, Held JP, Intering N, Kwakkel G, Pohl J, Reisman DS, Schwarz A, Sunnerhagen KS. Standardized international manual of the Fugl-Meyer assessment of motor function after stroke. Neurorehabilitation and neural repair. 2026 Apr;40(4):306-19. https://doi.org/10.1177/15459683251412300.
19. Vidmar T, Goljar Kregar N, Puh U. Reliability of the Modified Ashworth Scale after stroke for 13 muscle groups. Arch Phys Med Rehabil. 2023;104(10):1670-1677. https://doi.org/10.1016/j.apmr.2023.04.008.
20. Leszczak J, Pniak B, Baran J, Guzik A. Reliability of biometric devices for measuring hand grip and finger pinch strength in stroke patients over 50: a prospective observational study. Scientific Reports. 2025 Aug 9;15(1):29188. https://doi.org/10.1038/s41598-025-12712-1.
21. de Sire A, Marotta N, Prestifilippo E, Agostini F, Parente A, Tasselli A, Mezian K, Vecchio M, Longo UG, Ammendolia A. Efficacy of Constraint-Induced Movement Therapy and mirror therapy in improving upper limb motor function and dexterity in post-stroke hemiparetic patients: a randomized controlled trial. La Clinica Terapeutica. 2025 Nov 20;176(6). https://doi.org/10.7417/CT.2025.5288.
22. Tenberg S, Mueller S, Vogt L, Roth C, Happ K, Scherer M, Behringer M, Niederer D. Comparative effectiveness of upper limb exercise interventions in individuals with stroke: a network meta-analysis. Stroke. 2023 Jul;54(7):1839-53. https://doi.org/10.1161/STROKEAHA.123.043110.
23. Saikaley M, Pauli G, Sun H, Serra JR, Iruthayarajah J, Teasell R. Network meta-analysis of non-conventional therapies for improving upper limb motor impairment poststroke. Stroke. 2022 Dec;53(12):3717-27. https://doi.org/10.1161/STROKEAHA.122.040687.
24. Paul J, Yamin T, Rajalakshm G. Comparative Effect of Modified Constraint Induced Movement Therapy, Proprioceptive Training and Task-Oriented Training on Functions of Upper Extremity among Stroke Patient. https://doi.org/10.15621/ijphy/2023/v10i2/1326.
25. Hayward KS, Kramer SF, Dalton EJ, Hughes GR, Brodtmann A, Churilov L, Cloud G, Corbett D, Jolliffe L, Kaffenberger T, Rethnam V. Timing and dose of upper limb motor intervention after stroke: a systematic review. Stroke. 2021 Nov;52(11):3706-17. https://doi.org/10.1161/STROKEAHA.121.034348.
26. Zhang K, Ding L, Wang X, Zhuang J, Tong S, Jia J, Guo X. Evidence of mirror therapy for recruitment of ipsilateral motor pathways in stroke recovery: A resting fMRI study. Neurotherapeutics. 2024 Mar 1;21(2):e00320. https://doi.org/10.1016/j.neurot.2024.e00320.
27. Muñoz-Gómez E, Ingles M, Aguilar-Rodríguez M, Sempere-Rubio N, Mollà-Casanova S, Serra-Añó P. Effects of mirror therapy on spasticity and sensory impairment after stroke: Systematic review and meta-analysis. PM&R. 2023 Nov;15(11):1478-92. https://doi.org/10.1002/pmrj.12964.

Original Article

Allied Medical Research Journal
Vol:4 No.3: 13-21
© The Author (s) 2026

Incidence of Varicose Veins and Edema among Female Nurses in Islamabad: A Comparative Study

Adn Tahir Chaudry[1], Ameera Aamar[2], Muhammad Haroon[2], Mohammad Suleman[3]

[1]*Lecturer, Rawal Institute of Health Sciences, Islamabad, Pakistan*
[2]*Clinical Instructor, Rawal Institute of Health Sciences, Islamabad, Pakistan*
[3]*Clinical Physiotherapist, The Physio Studio, Islamabad, Pakistan*

ABSTRACT

Background: Varicose veins (VV) and edema are common chronic venous disorders affecting quality of life. Female nurses are at greater risk because of long hours standing and pregnancy, but incidence data available on these conditions in Pakistan is scarce. This study was aimed to compare the incidence of VV and edema among female nurses in Islamabad and to identify associated risk factors.

Methods: This descriptive, comparative, and cross-sectional study recruited 356 female nurses from government, semi-government, and private hospitals in Islamabad, selected through stratified random sampling from a pool of 4297 female nurses. Information was obtained using a validated questionnaire covering demographic data, disease status, family history, pregnancy, and risk factors, with edema graded on a pitting scale and VV classification based on the CEAP system. Descriptive statistics and Chi-square test ($p<0.05$) were used for analysis in SPSS version 22.

Results: Among 356 nurses, 51 (14.35%) were unaffected. Of the remainder, edema alone was found in 161 (52.8%), VV alone in 52 (17.0%), and both conditions in 92 (30.2%). Prolonged standing (35.1% and 18.7% respectively) was the most common risk factor for both edema and VV. Significant associations with disease occurrence were found for pregnancy ($\chi^2=12.046$, $p=0.017$) and family history ($\chi^2=19.123$, $p=0.004$), with 45.9% of affected nurses not receiving treatment.

Conclusion: VV was found to be less frequent than edema in female nurses of Islamabad. Family history, pregnancy, and prolonged standing were identified as major risk factors, underscoring the importance of screening and ergonomic modifications for nurses.

Keywords: *Edema, Nurses, Occupational Risk Factors, Pregnancy, Prevalence, Varicose Veins.*

Received: May 28th, 2026**; Revised:** June 26th, 2026**; Accepted:** July 10th, 2026

Corresponding Email: adntahir007@gmail.com

DOI: https://doi.org/10.59564/amrj/04.03/002

INTRODUCTION

Varicose veins (VV) are a common chronic venous disorder characterized by dilated, tortuous subcutaneous veins of the lower limbs and represent a substantial cause of morbidity that adversely affects patients' quality of life, occasionally with serious complications[1,2]. Left untreated, chronic venous disease may progress to venous ulceration and other manifestations of chronic venous insufficiency[1]. Guidelines from the European Society for Vascular Surgery similarly identify chronic venous disease as a common and progressive systemic disorder of the venous circulation[3]. Pregnancy has been identified as an important contributing risk factor for VV, given the comparatively low prevalence of the disease among non-pregnant women[4].

Edema, defined as the abnormal accumulation of fluid within body tissues, most commonly affects the legs and feet and may result from excessive fluid retention, poor venous drainage, prolonged standing, or vasodilating medications, as well as systemic causes such as heart failure, pulmonary hypertension, renal sodium retention, and pregnancy[5]. Atypical presentations, including pelvic congestion, are also recognized among the clinical spectrum of venous disease, arising

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

principally from valvular incompetence of the ovarian veins and associated venous reflux[6,7]. Edema occurring in this context is generally self-limited and predominantly affects pregnant and older women, and is managed with a combination of conservative and surgical approaches[8].

Prolonged standing at work has been linked to a range of adverse health outcomes, including low back and leg pain, cardiovascular strain, fatigue, and lower-limb edema, particularly among non-pregnant women[9]. Epidemiological estimates of VV prevalence vary considerably by population and geography. A global systematic review with pooled prevalence analysis reported VV (CEAP C2) in 19% of the adult population, with prevalence highest in Western Europe[10]. In Saudi Arabia, a prevalence of 42% has been reported among teachers, the majority of whom were female[11]. Among pregnant women, lower-limb VV have been reported in 20–40% of cases[12], with pregnancy reported to exacerbate pre-existing disease in around 11% of affected women and to precipitate initial symptoms in as many as 67%; the disease most frequently first becomes evident during the second trimester, reflecting the influence of hormonal change on venous tone[7].

Circulatory disorders continue to be one of the most important medical and social issues around the world due to their high prevalence, clinical impact, and repercussions on the quality of life[7]. The systemic nature of varicose disease means that its occurrence during pregnancy can complicate labor, delivery, and the postpartum period, contributing to increased maternal morbidity[13]. Lower-limb edema in this context, though also caused by lymphatic, traumatic, inflammatory, or malignant processes, is frequently one of the principal symptoms experienced by women in the third trimester as a consequence of venous insufficiency[8].

Occupational studies among nurses specifically have highlighted a disproportionately high burden of venous disease in this workforce[14]. A UK-based community nursing survey similarly found chronic edema to be common, with risk strongly associated with an aging, increasingly obese population and comorbidities such as diabetes mellitus, heart failure, and reduced mobility, although distinguishing primary from secondary chronic edema remained a limitation of that work[15]. Despite this recognized occupational vulnerability, comparative data quantifying and contrasting the incidence of VV and edema, and their associated risk factors among nurses in Pakistan are scarce. This study was therefore designed to compare the incidence of VV and edema among female nurses in Islamabad. The specific objectives were to determine and compare the incidence of the two conditions in this population, and to identify significant risk factors associated with their development.

METHODOLOGY

Study Design and Setting

A descriptive, comparative, cross-sectional study was conducted among female nurses employed in government, semi-government, and private hospitals across Islamabad, Pakistan. Data collection was carried out over six months, from May to October 2023, following approval of the research protocol. This study is reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cross-sectional studies.

Study Population and Sample Size

The sampling frame comprised 4,297 female nurses drawn from 22 government, semi-government, and private hospitals in Islamabad. Using the Raosoft sample size calculator (95% confidence level, 5% margin of error, 50% response distribution), a minimum sample of 353 was recommended; a final sample of 356 female nurses was recruited.

Sampling Technique

Stratified random sampling was employed. Hospitals were first stratified by sector (government/semi-government and private), and nurses were further sub-stratified by clinical work area (emergency, general, and private wards). Of the total sampling frame, 1,633 nurses (38%) were employed in government hospitals, and 2,664 (61%) in private/semi-government hospitals; sample allocation to each stratum was made proportionate to this distribution. Within each stratum, participants were selected using simple random sampling from a numbered list of eligible nurses.

Inclusion and Exclusion Criteria

Female nurses currently working in clinical positions in public or private hospitals in Islamabad were eligible for inclusion. Male nurses, nurses unwilling to participate, and

nurses reporting conditions other than VV or edema were excluded.

Data Collection Tool and Clinical Assessment

Data were collected using a self-structured, interviewer-administered questionnaire developed specifically for this study and organized into four domains: (i) demographic characteristics (age, educational qualification, height, weight, BMI, marital status); (ii) disease and family history, including current treatment-seeking status; (iii) pregnancy status and occupational exposure variables (duration of sitting and standing, obesity, age); and (iv) symptom burden, including leg pain over the preceding four weeks, rated using a validated 0–5 pain-rating (faces/numeric) scale. The content validity of the questionnaire was assessed using Pearson correlation, and internal consistency was tested using Cronbach's alpha after pilot testing with 15 participants before the full-scale data collection; items with poor item–total correlations were revised before the final administration.

The depth and recovery time of the skin depression after digital pressure were used to assess edema clinically, and were scored from 0+ (no pitting) to 4+ (severe, 8 mm depression not resolving within 2 minutes). Intermediate scores of 1+ (mild, 2 mm, rapidly resolving), 2+ (moderate, 4 mm, resolving in 10–15 seconds), and 3+ (moderately severe, 6 mm, persisting up to 1 minute) were also recorded. To aid objective comparison, calf and ankle girth were also measured at four fixed points (mid-thigh or widest calf, below calf, smallest ankle, and base of toes).

The VV were evaluated and graded according to internationally accepted classification system (the CEAP system), which was applied at the clinical (C) level in this study: C0 (no visible or palpable signs of venous disease), C1 (telangiectasia or reticular veins), C2 (varicose veins), C3 (Edema of venous origin), C4 (skin changes like pigmentation, eczema or atrophies blanche), C5 (healed venous leg ulcer), and C6 (active venous leg ulcer). In addition, each incidence of VV was categorized by recurrence (primary vs. recurrent), anatomic distribution (unilateral vs. bilateral), symptom status (symptomatic vs. asymptomatic), and presence of truncal venous reflux (present vs. absent), to provide a multidimensional clinical description of VV beyond prevalence alone.

Ethical Consideration

The study was reviewed and approved by the Ethical Review Committee (ERC Reference No: FMRL-IRB/2023/036) of Sarhad University, Peshawar, Pakistan. Informed consent was obtained from all participants before data collection, and participant information was kept confidential throughout the study.

Statistical Analysis

The data were coded and analyzed by SPSS version 22.0. All study variables were analyzed using descriptive statistics (frequencies and percentages). Associations between disease occurrence and pregnancy status and between disease occurrence and family history were evaluated using the Chi-square test, with p-values <0.05 considered statistically significant.

RESULTS

A total of 356 female nurses were enrolled in the study. Fifty-one participants (14.35%) reported neither VV nor edema; results are therefore reported for the remaining participants who had at least one of these conditions.

Demographic Characteristics

The average age of the respondents was 37.93±9.8 years. The largest proportion of nurses (33.1%) was aged 30–39 years, followed by 27.5% aged 20–29 years, 24.2% aged 40–49 years, and 14.75% aged 50–59 years. Nearly half (41.9%) held a Nursing Diploma, 37.3% held an FSc qualification, and 20.6% held a Matric qualification as their highest qualification. Regarding BMI, most respondents were overweight (68.2%), while 28.8% were of normal weight, 2.6% were obese, and 0.3% were underweight. The marital status of the majority of nurses (60.3%) was married, with 33.1% unmarried and 3.3% each widowed or divorced (Table 1).

Table 1: Demographic Characteristics of Respondents (n=356)		
Variable	**Frequency**	**Percentage (%)**
Age (years)		
20–29	84	27.5
30–39	101	33.1
40–49	74	24.2
50–59	45	14.75
Qualification		
FSc	114	37.3
Matric	63	20.6
Nursing Diploma	128	41.9
BMI		
Below 18.5 (Underweight)	1	0.3
18.5–24.9 (Normal)	88	28.8
25.0–39.9 (Overweight)	333*	10.8
40.0 & above (Obese)	8	2.6
Marital Status		
Married	184	60.3
Unmarried	101	33.1
Widowed	10	3.3
Divorced	10	3.3

Prevalence and Family History of Disease

Of the 305 nurses who were determined to have at least one venous condition, 52.8% had edema, 30.2% had both VV and edema, and 17.0% had VV only. As for family history, 24.6% reported a family history of edema alone, 22.0% reported having a family history of both conditions, 8.2% reported a family history of VV alone, and 45.2% reported no family history of either condition. Fewer than half of affected nurses (45.9%) were receiving no treatment for their disease, while 27.9% were being treated for edema, 15.4% for VV, and 10.5% for both. Only 16.7% of respondents reported being pregnant at the time of the study. The great majority of nurses (93.4%) reported leg pain within the previous four weeks, with pain severity most commonly rated as "hurts even more" (34.8%) or "hurts little more" (28.5%) on the pain-rating scale (Table 2).

Table 2: Prevalence, Family History, Treatment Status, and Pregnancy (n=305)		
Variable	**Frequency**	**Percentage (%)**
Disease Found		
Varicose Vein	52	17.0
Edema	161	52.8
Both	92	30.2
Family History		
Varicose Vein	25	8.2
Edema	75	24.6
Both	67	22.0
None	138	45.2
Currently Receiving Treatment For		
Varicose Vein	47	15.4

Edema	85	27.9
Both	32	10.5
None	140	45.9
Currently Pregnant		
Yes	51	16.7
No	253	83.0
Leg Pain (previous 4 weeks)		
Yes	285	93.4
No	20	6.6
Pain Severity		
Hurts a little bit	21	6.9
Hurts a little more	87	28.5
Hurts even more	106	34.8
Hurts a whole lot	64	21.0
Hurts worst	6	2.0

Risk Factors and Pitting Edema Severity

Prolonged standing was the most commonly reported risk factor for edema (35.1%), followed by prolonged sitting (20.7%), pregnancy (11.8%), age (8.2%), and obesity (6.6%). Among nurses with edema, moderate pitting edema (2+) was most common (33.4%), followed closely by moderately severe pitting (3+) (31.5%); mild (1+) and severe (4+) pitting were seen in 14.1% and 2.6% of cases, respectively (Table 3).

Table 3: Risk Factors for Edema and Pitting Edema Grade

Variable	Frequency	Percentage (%)
Risk Factors for Edema		
Obesity	20	6.6
Age	25	8.2
Prolonged sitting	63	20.7
Prolonged standing	107	35.1
Pregnancy	36	11.8
Pitting Edema Grade		
Mild (1+)	43	14.1
Moderate (2+)	102	33.4
Moderately Severe (3+)	96	31.5
Severe (4+)	8	2.6

Clinical Variables of Varicose Veins

Among nurses with VV, prolonged standing was again the most frequently cited risk factor (18.7%), followed by prolonged sitting (10.2%), obesity (8.5%), age (7.5%), and pregnancy (5.9%). Disease was primary in 28.5% and recurrent in 23.3% of cases; involvement was unilateral in 27.2% and bilateral in 25.2%.

Symptomatic disease (22.6%) was less common than asymptomatic disease (29.8%), and truncal reflux was demonstrated in 16.7% of affected nurses. On CEAP clinical grading, varicose veins (C2) and edema (C3) were each identified in 18.0% of cases, telangiectasia (C1) in 12.5%, eczema/atrophic blanche (C4) in 2.3%, and a healed venous leg ulcer (C5) in 0.3% (Table 4).

Table 4: Clinical Variables of Varicose Veins		
Variable	**Frequency**	**Percentage (%)**
Risk Factors for Varicose Veins		
Obesity	26	8.5
Age	23	7.5
Prolonged sitting	31	10.2
Prolonged standing	57	18.7
Pregnancy	18	5.9
Recurrence		
Primary	87	28.5
Recurrent	71	23.3
Location		
Unilateral	83	27.2
Bilateral	77	25.2
Symptom Status		
Symptomatic	69	22.6
Asymptomatic	91	29.8
Reflux		
Truncal reflux	51	16.7
No truncal reflux	109	35.7
CEAP Clinical Grade		
C1 – Telangiectases	38	12.5
C2 – Varicose veins	55	18.0
C3 – Edema	55	18.0
C4 – Eczema/Atrophie blanche	7	2.3
C5 – Healed venous leg ulcer	1	0.3

Association of Disease with Pregnancy and Family History

A statistically significant association was found between pregnancy status and disease occurrence (χ^2=12.046, df=4, p=0.017). Among the 51 pregnant nurses, 2 had varicose veins, 31 had edema, and 18 had both conditions; among the 253 non-pregnant nurses, 50 had varicose veins, 130 had edema, and 74 had both conditions (Table 5).

Table 5: Relationship of Pregnancy with Disease Occurrence				
Pregnancy Status	**Varicose Vein**	**Edema**	**Both**	**Total**
Pregnant	2	31	18	51
Not Pregnant	50	130	74	253
Total	52	161	92	305

Pearson Chi-Square = 12.046, df = 4, p = 0.017 (significant)

A significant association was also observed between family history and disease occurrence (χ^2=19.123, df=6, p=0.004). Among nurses with a family history of VV (n=25), 6 had VV, 7 had edema, and 12 had both; among those with a family history of edema (n=75), 4 had VV, 50 had edema, and 21 had both; among those with a family history of both conditions (n=67), 16 had VV, 29 had edema, and 22 had both (Table 6).

Table 6: Relationship of Family History with Disease Occurrence

Family History	Varicose Vein	Edema	Both	Total
Varicose Vein	6	7	12	25
Edema	4	50	21	75
Both	16	29	22	67
None	26	75	37	138
Total	52	161	92	305

Pearson Chi-Square = 19.123, df = 6, p = 0.004 (significant)

DISCUSSION

This comparative cross-sectional study found edema to be considerably more common than VV among female nurses in Islamabad (52.8% vs 17.0%, with a further 30.2% presenting both conditions concurrently), against an overall background prevalence of at least one condition in 85.65% of the sample. This distribution is similar to the overall literature reporting that chronic venous disease is highly variable, with a range of pooled estimates of VV and telangiectasia across different populations and diagnostic criteria[3]. Occupational surveys among nurses elsewhere have similarly documented a substantial burden of venous disease: a study in Mosul, Iraq, reported a VV incidence of 20% across occupational groups, with a comparable proportion of affected men (51%) and women (50%)[16], while a clinical study of lower-limb VV reported a peak incidence of 62% among nurses aged 21–40 years, with limb heaviness the most common presenting symptom[17].

The proportion of respondents with a family history of either condition (54.8% combined) and the statistically significant association demonstrated between family history and disease occurrence (p=0.004) support a hereditary predisposition to venous disease, echoing UK community-based data in which chronic Edema was similarly found to cluster with modifiable and non-modifiable comorbid risk factors, including diabetes, heart failure, obesity, and reduced mobility[14]. A global meta-analysis of healthcare workers, alongside occupational surveys of nurses in Saudi Arabia, Egypt, and Nepal, has likewise identified family history, alongside age, marital status, and physically demanding duties, as significant independent predictors of VV[15,18,19,20]. Prolonged standing emerged as the dominant risk factor for both Edema (35.1%) and VV (18.7%) in this cohort, consistent with occupational health evidence linking sustained standing posture to lower-limb pain, fatigue, and edema[9]. A cross-sectional study of nurses at Dhulikhel Hospital similarly found VV in 83% of participants, with each additional hour of daily standing increasing the odds of disease by 27-fold in adjusted analysis[19]. Physiological studies of prolonged static occupational posture provide a mechanistic explanation for this association, having demonstrated measurable increases in leg volume and venous pooling with sustained standing, attributable to hydrostatic pressure effects on venous return[21].

Pregnancy was reported by 16.7% of participants. It was significantly associated with disease occurrence (p=0.017), with pregnant nurses disproportionately affected by Edema (31/51) and by concurrent VV and Edema (18/51) relative to VV alone (2/51). This pattern mirrors evidence that lower-limb Edema, more than isolated VV, is a prominent feature of the third trimester as a consequence of venous insufficiency and mechanical compression from the gravid uterus,8 and aligns with reports that VV risk rises with each successive pregnancy, reaching as high as 37–38% by the third and fourth pregnancies[7]. A large multicenter analysis has similarly identified specific pregnancy-related factors, including multiparity and pregnancy complications, as significant contributors to the development of VV[1].

Only about half of the affected nurses in this study (54.1%) were receiving any treatment for their condition, a finding that raises concern about under-recognition and under-treatment of venous disease within this occupational group, consistent with broader observations that chronic Edema remains an under-recognized and undertreated health problem globally[22]. In addition, high prevalence of self-reported leg pain (93.4%) and moderate to moderately severe pitting edema (64.9% combined) further highlight the

symptomatic burden of affected nurses' lives, and confirm that greater symptom burden in venous disease is associated with poorer HRQoL23[23].

The clinical profile of VV observed in this study with the majority of cases falling within CEAP grades C1–C3 and truncal reflux demonstrated in only 16.7% of cases suggests that most disease in this cohort was at an early-to-moderate stage, offering a window for conservative and preventive intervention before progression to more severe manifestations such as skin changes or ulceration[24]. Conservative measures, including structured patient education and hydrotherapy, have been shown to meaningfully improve both knowledge and symptom burden among patients with lower-limb VV[25], while compression therapy and activity-based approaches remain central to the non-surgical management of chronic edema[26].

This study has some limitations that should be considered when interpreting its findings. Disease status was determined through a structured questionnaire and clinical grading rather than objective duplex or Doppler ultrasonographic confirmation, which may have introduced misclassification bias, and the cross-sectional design precludes causal inference regarding the identified risk factors; the study was also restricted to VV and edema, and published comparative data on edema incidence among nurses in Pakistan remain limited, constraining direct benchmarking of the present findings. Based on these findings, ultrasonographic confirmation of disease status is recommended for future studies, along with longitudinal designs to establish causal relationships. From a practice standpoint, activity-oriented anti-edema therapeutic approaches, graduated compression hosiery, and endovenous thermal ablation (EVLA/RFA) for symptomatic VV warrant consideration, alongside simple preventive measures such as avoiding prolonged standing or sitting, maintaining a healthy body weight, regular position changes and lower-limb exercise during shifts, and leg elevation during rest, particularly among nurses with a positive family history or who are pregnant.

CONCLUSION

The findings revealed that edema affected a significantly larger number of female nurses than VV, with almost one-third of the affected nurses having both edema and VV. Both family history and pregnancy were significant factors in the occurrence of both diseases, and prolonged standing was found to be the most important modifiable occupational risk factor for both diseases. Many of those affected were not receiving treatment, indicating a lack of treatment-seeking behavior in this occupation. Based on the findings of this study, targeted screening, ergonomic adjustments, and preventive health education programs could be implemented for nurses, particularly for pregnant women and those with a positive family history.

Author Contributions

ATC: Conception and Design, Data Collection, and Supervision.
AA: Study Design, Data Analysis, and Interpretation**.**
MH: Literature Review, Manuscript Writing, and Critical Review.
MS: Data Collection, Manuscript Editing, and Final Approval.

Funding Source

No sources.

Conflict of Interest

The authors declared no conflict of interest.

Data Availability

Data will be available on request.

Ethical Approval

The study was reviewed and approved by the Ethical Review Committee (ERC Reference No: FMRL-IRB/2023/036) of Sarhad University, Peshawar, Pakistan.

REFERENCES

1. DeCarlo C, Boitano LT, Waller HD, Pendleton AA, Latz CA, Tanious A, Kim Y, Mohapatra A, Dua A. Pregnancy conditions and complications associated with the development of varicose veins. Journal of Vascular Surgery: Venous and Lymphatic Disorders. 2022 Jul 1;10(4):872-8. doi: https://doi.org/10.1016/j.jvsv.2022.01.003
2. Gentile G, Tambuzzi S, Boracchi M, Gobbo AD, Bailo P, Zoia R. Fatal hemorrhage from peripheral varicose vein rupture. Autopsy and Case Reports. 2021 Sep 24;11:e2021330. doi: https://doi.org/10.4322/acr.2021.330
3. De Maeseneer MG, Kakkos SK, Aherne T, Baekgaard N, Black S, Blomgren L, Giannoukas A, Gohel M, de Graaf R, Hamel-Desnos C, Jawien A. Editor's choice–European Society for Vascular Surgery (ESVS) 2022 clinical practice guidelines on the management of chronic venous disease of the lower limbs. European Journal of Vascular and Endovascular Surgery. 2022 Feb 1;63(2):184-267. doi: https://doi.org/10.1016/j.ejvs.2021.12.024
4. Zhang J, Sun JL. Severe venous thromboembolism in the puerperal period caused by thrombosis: A case report. World Journal of Clinical Cases. 2020 Apr 6;8(7):1311. doi: https://doi.org/10.12998/wjcc.v8.i7.1311

5. Patel H, Skok CC, DeMarco A. Peripheral edema: evaluation and management in primary care. American Family Physician. 2022 Nov;106(5):557-64.
6. Fukaya E, Kolluri R. Nonsurgical management of chronic venous insufficiency. New England Journal of Medicine. 2024 Dec 19;391(24):2350-9. doi: https://doi.org/10.1056/NEJMcp2310224
7. Karaś R, Lepich T, Spyra R, Bajor G. Varicose veins of the lower limbs in pregnant women. Pielęgniarstwo Chirurgiczne i Angiologiczne/Surgical and Vascular Nursing. 2024;18(3):75-80. doi: https://doi.org/10.5114/.2024.144113
8. Aslam MR, Muhammad Asif H, Ahmad K, Jabbar S, Hayee A, Sagheer MS, Rehman JU, Khalid S, Hashmi AS, Rajpoot SR, Sharif A. Global impact and contributing factors in varicose vein disease development. SAGE Open Medicine. 2022 Aug;10:20503121221118992. doi: https://doi.org/10.1177/20503121221118992
9. Hirsch T, Wahl U, Rabe E. Venous disorders as an occupational disease–a systematic review on epidemiology, pathophysiology, and modification strategies. Vasa. 2024 Mar 27. doi: https://doi.org/10.1024/0301-1526/a001116
10. Salim S, Machin M, Patterson BO, Onida S, Davies AH. Global epidemiology of chronic venous disease: a systematic review with pooled prevalence analysis. Annals of surgery. 2021 Dec 1;274(6):971-6. doi: https://doi.org/10.1097/SLA.0000000000004631
11. Dalboh A, Alshehri NA, Alrafie AA, Bakri KA. Prevalence and awareness of varicose veins among teachers in Abha, Saudi Arabia. Journal of Family Medicine and Primary Care. 2020 Sep 1;9(9):4784-7. doi: https://doi.org/10.4103/jfmpc.jfmpc_490_20.
12. Youn YJ, Lee J. Chronic venous insufficiency and varicose veins of the lower extremities. The Korean journal of internal medicine. 2018 Oct 26;34(2):269. doi: https://doi.org/10.3904/kjim.2018.230
13. Nam JY, Park EC. The relationship between severe maternal morbidity and a risk of postpartum readmission among Korean women: a nationwide population-based cohort study. BMC Pregnancy and Childbirth. 2020 Mar 6;20(1):148. doi: https://doi.org/10.1186/s12884-020-2820-7
14. He QF, Cai JY, Cheng M, Feng SJ, Lu QY, Wang F. Global prevalence and risk factors of varicose veins among health care workers: a systematic review and meta-analysis. BMC nursing. 2025 May 16;24(1):550. doi: https://doi.org/10.1186/s12912-025-03155-0
15. Moffatt CJ, Gaskin R, Sykorova M, Dring E, Aubeeluck A, Franks PJ, Windrum P, Mercier G, Pinnington L, Quéré I. Prevalence and risk factors for chronic edema in UK community nursing services. Lymphatic Research and Biology. 2019 Apr 1;17(2):147-54. doi: https://doi.org/10.1089/lrb.2018.0086
16. Mahmood HJ, Ibrahim RH, Abdulghani MF, Al-Zaidy ZF. A quantitative study on evidence of varicose veins among different occupational backgrounds in the city of Mosul Iraq. Malaysian Journal of Public Health Medicine. 2021 Dec 28;21(3):36-41. doi: https://doi.org/10.37268/mjphm/vol.21/no.3/art.1105
17. Hajar A. Machine Learning Reveals: Yoga Practice Significantly Reduces Varicose Vein Severity. Journal of Artificial Intelligence and System Modelling. 2025 Sep 30;3(03):84-104. https://doi.org/10.22034/jaism.2025.537892.1147
18. AlBader B, Sallam A, Moukaddem A, Alanazi K, Almohammed S, Aldabas H, Algmaizi S. Prevalence of varicose veins among nurses at different departments in a single tertiary care center in Riyadh. Cureus. 2020 Dec 27;12(12). doi: https://doi.org/10.7759/cureus.12319
19. Shakya R, Karmacharya RM, Shrestha R, Shrestha A. Varicose veins and its risk factors among nurses at Dhulikhel hospital: a cross sectional study. BMC nursing. 2020 Feb 3;19(1):8. doi: 10.1186/s12912-020-0401-8
20. Abou-ElWafa HS, El-Metwaly AA, El-Gilany AH. Lower limb varicose veins among nurses: a single center cross-sectional study in Mansoura, Egypt. Indian journal of occupational and environmental medicine. 2020 Sep 1;24(3):172-7. doi: https://doi.org/10.4103/ijoem.IJOEM_264_19
21. Wall R, Garcia G, Läubli T, Seibt R, Rieger MA, Martin B, Steinhilber B. Physiological changes during prolonged standing and walking considering age, gender and standing work experience. Ergonomics. 2020 May 3;63(5):579-92. doi: https://doi.org/10.1080/00140139.2020.1725145
22. Keast DH, Moffatt C, Janmohammad A. Lymphedema impact and prevalence international study: the Canadian data. Lymphatic Research and Biology. 2019 Apr 1;17(2):178-86. doi: https://doi.org/10.1089/lrb.2019.0014
23. Maisto M, Marzocchi A, Piccolo V, Ciampaglia R, De Vivo M, Tenore GC. A Multi-Component Nutraceutical Formulation for the Management of Vascular and Inflammatory Alterations Characteristic of the Oedema Disorders. Nutrients. 2026 Feb 4;18(3):523. doi: https://doi.org/10.3390/nu18030523
24. Chung JH, Heo S. Varicose veins and the diagnosis of chronic venous disease in the lower extremities. Journal of chest surgery. 2023 Nov 23;57(2):109. doi: https://doi.org/10.5090/jcs.23.110
25. Wang C, Sun H, Wang F, Shen S, Li J. Nursing Care Strategies for Varicose Veins: A Narrative Review. Journal of Multidisciplinary Healthcare. 2025 Dec 31:7667-78. doi: https://doi.org/10.2147/JMDH.S564809
26. Thibert A, Briche N, Vernizeau BD, Mougin-Guillaume F, Beliard S. Systematic review of adapted physical activity and therapeutic education of patients with chronic venous disease. Journal of Vascular Surgery: Venous and Lymphatic Disorders. 2022 Nov 1;10(6):1385-400. doi: https://doi.org/10.1016/j.jvsv.2022.05.004

Original Article

Allied Medical Research Journal

Vol:4 No.3: 22-30

© The Author (s) 2026

The Efficacy of Instrument-Assisted Soft Tissue Mobilization as an Adjunct to Therapeutic Exercise in Knee Osteoarthritis: A Parallel-Group Randomized Clinical Trial

Faiza Kalsoom[1], Mehreen Ali[2], Khakshan Taveel[3], Rabbia Irshad[2], , Zona Mehreen[4]

[1]*Mukabbir University of Science and Technology, Gujrat, Pakistan*
[2]*Malik Haider Hospital, Gujrat, Pakistan*
[3]*Physio Care Centre, Gujrat, Pakistan*
[4]*SK Physio & Aesthetic Clinic, Islamabad, Pakistan*

ABSTRACT

Background: In adults over 40, knee osteoarthritis (OA) commonly causes pain and dysfunction, often associated with tight soft tissues, quadriceps weakness, and limited range of motion. While therapeutic exercise is the main non-surgical treatment, its effectiveness can be limited by myofascial restriction. This study aimed to assess the effectiveness and safety of Instrument-Assisted Soft Tissue Mobilization (IASTM) combined with the exercise program compared with exercise alone in reducing pain, knee range of motion, and function in individuals with knee OA.

Methods: This is a parallel-group randomized trial with participant blinding, conducted over 12 months at two centers: Physio Care Center, Gujarat, and SK Physio OPD, Islamabad. Fifty patients with knee OA (Kellgren-Lawrence grade II-III) were randomly assigned to Group A (exercise + IASTM) or Group B (exercise only), with 25 patients in each group. All patients received 3 sessions per week for 6 weeks and were assessed at baseline and at 6 weeks using the Visual Analog Scale, Goniometry, and WOMAC index respectively.

Results: Both groups demonstrated improvement, with Group A (exercise + IASTM) demonstrating more pain reduction, knee flexion improvement, and residual effects compared to Group B (exercise only). All results in Group A were $p<0.01$ compared with Group B, which reduced extension deficits and showed better WOMAC scores.

Conclusion: The findings revealed that combining IASTM with therapeutic exercise yielded greater improvements in pain, knee mobility, and function than exercise alone in knee OA, making it a valuable complementary tool for traditional knee rehabilitation.

Keywords: *Exercise; Knee; Mobility; Osteoarthritis; Range of Motion.*

Received: May 4th, 2026; **Revised:** June 17th, 2026; **Accepted:** July 2nd, 2026

Corresponding Email: faiza.kalsoom@mukabbiruniversity.edu.pk

DOI: https://doi.org/10.59564/amrj/04.03/003

INTRODUCTION

Osteoarthritis (OA) of the knee is a leading cause of persistent pain, disability, and reduced quality of life in adults over 401. According to the 2021 Global Burden of Disease, the number of people with knee OA exceeds 374 million, and its prevalence is projected to rise further until 2035 as the global population ages and expands[1]. This trajectory can further increase costs of care and productivity losses, leading to increasing significance of conservative treatments that are effective, affordable, and easily replicable[1].

Knee OA does not only involve loss of cartilage. The overall presentation is driven by synovium inflammation, sub-chondral bone remodeling, quadriceps weakness, and tightness of the fascia or myofascial tissue around the knee[2]. These alterations in soft tissues can perpetuate poor movement patterns, restrict joint mobility, and reduce the benefits of exercise training[3]. This has led to an increased focus on rehabilitation strategies that address both the joint itself and the surrounding soft tissues.

This has led to increased interest in manual therapy techniques that directly address fascial restrictions to enhance outcomes from exercise-based therapy, especially when some stiffness and decreased mobility persist after routine care[3,4]. Instrument-Assisted Soft Tissue Mobilization (IASTM) is one such manual approach, in which purpose-built tools are used to deliver controlled mechanical force to areas of soft tissue restriction. Proposed mechanisms

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

include greater tissue extensibility, release of fascial adhesions, and activation of local fibroblasts[5]. A growing set of trials has started evaluating IASTM specifically in people with knee OA, with reported benefits spanning gait-related muscle activity when paired with exercise,[5] improved clinical performance after iliotibial band treatment,[6] better gait kinetics,[7] gains in lower-limb strength and sit-to-stand ability,[8] and wider improvements in function and quality of life,[9] and a meta-analysis spanning several musculoskeletal conditions concluded that IASTM produces a meaningful improvement in range of motion[10]. Other early-stage work points the same way, describing symptom relief among knee OA patients who received IASTM[9,10]. Thus, the existing literature varies considerably in treatment dose, comparison arms, and outcome measures, and relatively few studies have tested IASTM layered onto a standardized exercise protocol against exercise by itself within a tightly controlled, parallel-group design with concealed allocation. It is therefore unclear whether IASTM provides any additional efficacy beyond exercise therapy in patients with radiographically confirmed knee OA (Kellgren-Lawrence grade II-III)[11]. To address this disconnect, the present parallel-group randomized controlled trial was designed to examine the effects of IASTM, in addition to therapeutic exercise, on knee pain (VAS), ROM (goniometry), and functional performance (WOMAC) in individuals with knee OA, to generate evidence that can be directly applied to clinical practice.

METHODOLOGY

Study Design and Setting

The trial was a single-masked, two-arm, parallel-group Randomized Controlled Trial conducted at two sites, namely Physio Care Center, Gujarat, and SK Physio OPD, Islamabad, from January to December 2025. Patients with knee osteoarthritis (Kellgren-Lawrence II-III) were randomly allocated to two groups in equal numbers, with concealment achieved using sequentially numbered, sealed, opaque envelopes.

Sample Size

Sample size was determined with G*Power (version 3.1.9.4), drawing on figures from an earlier RCT titled “A comparative study of the efficacy of instrument-assisted soft tissue mobilization and massage techniques in patients with patellofemoral joint pain,”[12] which had reported post-treatment Numeric/Visual Analogue Scale (VAS) values of 2.25 ± 1.38 in the IASTM arm versus 3.58 ± 1.38 in the Tui-na arm. Based on these figures, the calculation indicated a minimum of roughly 17 participants per group would be needed. To guard against a type II error, allow for possible dropouts, and limit selection and attrition bias, the target was raised to 50 participants overall, with 25 randomly assigned to each of Group A and Group B.

Selection Criteria

Inclusion criteria: Participants were included if they (i) had a confirmed diagnosis of knee osteoarthritis (Kellgren-Lawrence grade II-III); (ii) were aged between 40 and 60 years; and (iii) were of either gender.

Exclusion criteria: Participants were excluded if they (i) did not meet the inclusion criteria, or (ii) had undergone recent lower-limb surgery.

Data Collection Procedure

Baseline testing took place in the week before treatment began. A trained assessor who did not know which group each participant belonged to recorded every outcome measure at baseline (Week 0) and again after treatment (Week 6).

- ***Visual Analogue Scale (VAS):*** Pain was rated on an 11-point scale from 0, meaning no pain, to 10, the worst pain imaginable, captured both while resting and during weight-bearing tasks such as climbing stairs; higher numbers reflected worse pain[13].

- ***Knee Range of Motion (Goniometry):*** This was measured by a standard universal goniometer at baseline and at six weeks. For extension, they were lying on their backs; for flexion, they were seated or lying on their stomachs with the knee free to hang. The fulcrum of the goniometer was placed over the lateral femoral epicondyle, the fixed arm over the greater trochanter, and the moving arm over the lateral malleolus. The knee should bend to 135°-140° and straighten completely to 0° (some individuals may have a small hyperextension of 5°-10°, which is normal for them). For mobility closer to these parameters, readings were considered good, whereas decreased flexion or inability to reach 0° extension indicated capsular or soft-tissue tightness[14].

- ***Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC):*** Function was assessed using

the self-report questionnaire, the WOMAC index, which was validated specifically for people with OA. It includes 24 items divided into 3 domains: pain (5 items – walking, using stairs, resting in bed, sitting or lying, and standing); stiffness (2 items – morning stiffness and stiffness later in the day); physical function (17 items – general day-to-day activities including walking, getting up from sitting, standing, bending, getting in and out of a car, shopping, dressing the feet, getting in and out of bed, bathing, sitting, using the toilet, and household activities of varying intensity). For each item, a Likert scale is used, ranging from "none" (0) to "extreme" (4), with a possible range of 0 to 96; higher scores indicate greater pain, stiffness, and disability, and lower scores indicate greater function[15].

Intervention Protocol

The duration of both study arms was six weeks, and the therapy sessions were 30 minutes in length, three times a week, on non-consecutive days (e.g., Monday, Wednesday, and Friday), supervised directly by a qualified physical therapist.

Group A: IASTM Combined with a Standardized Therapeutic Exercise Protocol

Each session opened with IASTM, delivered with a stainless-steel tool over the soft tissue surrounding the affected knee — the quadriceps (rectus femoris and vastus medialis/lateralis), iliotibial band, hamstrings, and peripatellar retinacular tissue. The therapist held the tool at roughly a 30°-60° angle against lightly lubricated skin, working proximal to distal with multi-directional sweeping strokes at a comfortable pressure for about 6-8 minutes, gradually increasing intensity over the six weeks based on how the tissue responded and the participant's pain tolerance. Exercise followed immediately after IASTM, using the same standardized protocol described below:

- **Quadriceps isometric/setting exercises, 3 sets of 10 repetitions, 5-10 second hold**
- **Straight leg raises, 3 sets of 10 repetitions**
- **Active-assisted knee flexion/extension range of motion exercises, 2 sets of 10 repetitions**
- **Hamstring stretching, 3 repetitions of 30-second static holds**
- **Terminal knee extension exercises using a resistance band, 3 sets of 10 repetitions**
- **Mini-squats/functional strengthening (as tolerated), 2 sets of 10 repetitions**

Group B: Therapeutic Exercise Protocol Alone

Group B followed the same exercise protocol (quadriceps setting, straight leg raises, active-assisted ROM work, hamstring stretching, terminal knee extension, and functional strengthening) at matching frequency, session length, and intensity progression, but skipped the IASTM step, filling the full 30 minutes with exercise alone, three times a week for six weeks.

Statistical Analysis

Analysis was carried out in SPSS version 25 (IBM Corp., USA). Normality of the data was checked with the Shapiro-Wilk test. Continuous variables are summarized as mean ± standard deviation, and categorical variables as counts and percentages. Paired-samples t-tests were used for within-group comparisons and independent-samples t-tests for between-group comparisons, with statistical significance set at $p < 0.05$.

Ethical Consideration

The approval was granted by the Institutional Ethical Review Committee at the Foundation of Medical Research and Lab (FMRL-IRB/2023/026) before the trial began, and every participant gave written informed consent prior to enrollment. The study followed Declaration of Helsinki principles and is registered on ClinicalTrials.gov under NCT07710209.

RESULTS

Fifty people with knee osteoarthritis were enrolled and split evenly into Group A (IASTM plus standardized exercise) and Group B (exercise only), 25 in each. No one dropped out over the six weeks, so all 50 completed both the baseline and post-intervention testing.

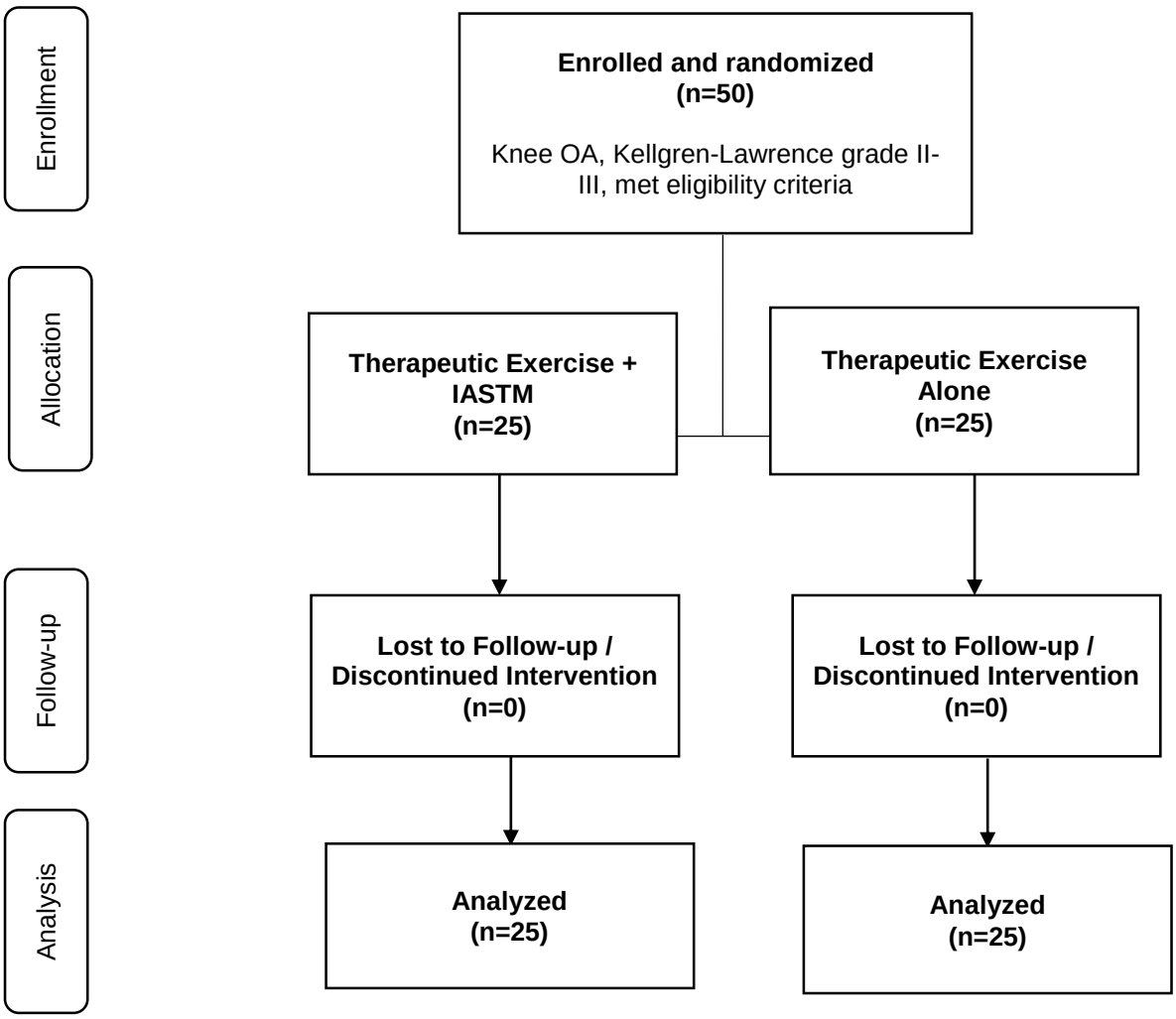

Figure 1. CONSORT flow diagram of participant enrollment, allocation, follow-up, and analysis

Each outcome variable was tested for normality using the Shapiro-Wilk test before running the main analyses. Most variables did ($p > 0.05$), both at baseline (WOMAC score, $p = 0.233$) and at Week 6 (WOMAC score, $p = 0.889$; knee flexion ROM, $p = 0.245$), with W-statistics of 0.949 and 0.980, respectively. With this in mind, throughout the data analysis, a parametric test, paired-samples t-test within groups and independent-samples t-test between groups, were used as planned.

Baseline demographics and clinical measures were compared across groups to check that randomization had worked as intended. Mean age was 48.44 ± 6.45 years in Group A versus 49.76 ± 6.31 years in Group B, a non-significant difference ($p = 0.468$), and the male-to-female split was similar between groups (10/15 in A, 11/14 in B; $p = 1.000$). None of the baseline outcome measures differed significantly either: VAS pain (7.16 ± 0.80 vs 6.84 ± 0.80, $p = 0.164$), knee flexion ROM (101.40° ± 6.24 vs 100.92° ± 6.98, $p = 0.799$), knee extension deficit (10.76° ± 3.02 vs 10.52° ± 3.20, $p = 0.786$), and WOMAC score (66.88 ± 7.79 vs 68.32 ± 7.30, $p = 0.503$), all $p > 0.05$ (Table 1). This supports attributing the post-intervention differences to treatment rather than to any starting imbalance between groups.

Group A's VAS score dropped from 7.16 ± 0.80 at baseline to 2.92 ± 1.15 by Week 6 ($p < 0.001$), while Group B's fell from 6.84 ± 0.80 to 4.12 ± 1.20 over the same period ($p < 0.001$). Comparing the two groups directly at Week 6, Group A's score was significantly lower than Group B's (2.92 ± 1.15 vs. 4.12 ± 1.20; $p < 0.001$) that is, pain relief was greater with IASTM added to exercise than with exercise on its own (Table 2).

Table 1. Baseline Demographic and Clinical Characteristics of Participants

Variable	Group A: IASTM + Exercise (n=25)	Group B: Exercise Only (n=25)	p-value
Age (years)	48.44 ± 6.45	49.76 ± 6.31	0.468
Gender (Male/Female)	10 / 15	11 / 14	1.000
VAS Baseline	7.16 ± 0.80	6.84 ± 0.80	0.164
Knee Flexion ROM Baseline (°)	101.40 ± 6.24	100.92 ± 6.98	0.799
Knee Extension Deficit Baseline (°)	10.76 ± 3.02	10.52 ± 3.20	0.786
WOMAC Baseline	66.88 ± 7.79	68.32 ± 7.30	0.503

Values are presented as Mean ± SD or n. Between-group comparison by independent-samples t-test (gender by chi-square test). Significance set at p < 0.05

Table 2. Within- and Between-Group Comparison of Pain (Visual Analogue Scale)

Parameter	Group A Baseline	Group A Week 6	Within-Group p	Group B Baseline	Group B Week 6	Within-Group p	Between-Group p (Week 6)
VAS	7.16 ± 0.80	2.92 ± 1.15	<0.001	6.84 ± 0.80	4.12 ± 1.20	<0.001	<0.001

Data are expressed as Mean ± SD. Paired-samples t-test for within-group comparison and independent-samples t-test for between-group comparison. Significance set at p < 0.05.

Both groups gained knee flexion range: Group A moved from 101.40° ± 6.24 to 123.88° ± 6.92 (p < 0.001), and Group B from 100.92° ± 6.98 to 113.44° ± 7.48 (p < 0.001). Extension deficit shrank in both arms too, from 10.76° ± 3.02 to 4.12° ± 2.85 in Group A (p < 0.001) and from 10.52° ± 3.20 to 7.16° ± 3.75 in Group B (p < 0.001). At Week 6, Group A's flexion ROM was significantly greater than Group B's (123.88° ± 6.92 vs 113.44° ± 7.48; p < 0.001), and its residual extension deficit was significantly smaller (4.12° ± 2.85 vs 7.16° ± 3.75; p = 0.002), both findings favor the IASTM group (Table 3).

Table 3. Within- and Between-Group Comparison of Knee Range of Motion

Parameter	Group A Baseline	Group A Week 6	Within-Group p	Group B Baseline	Group B Week 6	Within-Group p	Between-Group p (Week 6)
Flexion ROM (°)	101.40 ± 6.24	123.88 ± 6.92	<0.001	100.92 ± 6.98	113.44 ± 7.48	<0.001	<0.001
Extension Deficit (°)	10.76 ± 3.02	4.12 ± 2.85	<0.001	10.52 ± 3.20	7.16 ± 3.75	<0.001	0.002

Values are presented as Mean ± SD. Within-group comparison by paired-samples t-test; between-group comparison by independent-samples t-test. Significance set at p < 0.05.

WOMAC scores also fell in both groups, Group A from 66.88 ± 7.79 to 33.96 ± 9.13 (p < 0.001), and Group B from 68.32 ± 7.30 to 50.32 ± 8.13 (p < 0.001). At Week 6, Group A's score was significantly better (lower) than Group B's (33.96 ± 9.13 vs 50.32 ± 8.13; p < 0.001), meaning the IASTM-plus-exercise combination produced a larger functional improvement than exercise alone (Table 4).

Table 4. Within- and Between-Group Comparison of Functional Status (WOMAC)

Parameter	Group A Baseline	Group A Week 6	Within-Group p	Group B Baseline	Group B Week 6	Within-Group p	Between-Group p (Week 6)
WOMAC Score	66.88 ± 7.79	33.96 ± 9.13	<0.001	68.32 ± 7.30	50.32 ± 8.13	<0.001	<0.001

Values are presented as Mean ± SD. Within-group comparison by paired-samples t-test; between-group comparison by independent-samples t-test. Significance set at p < 0.05.

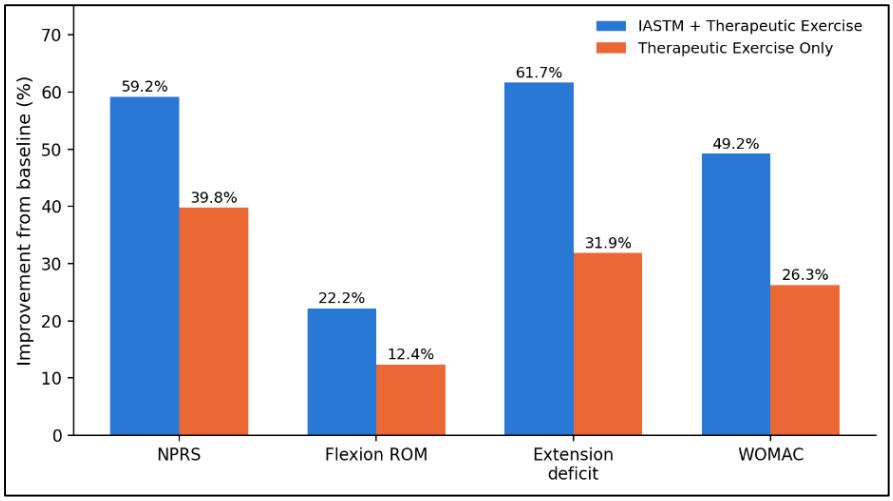

Figure 2. Percentage improvement in outcome measures at Week 6 by group

Figure 2 shows the percent change from baseline for each outcome measure in both groups at Week 6. Group A's percentage improvements were greater than Group B's on all four parameters (pain (VAS), knee flexion ROM, extension deficit, and WOMAC score) and greatest on extension deficit (61.7% vs 31.9%) and pain (59.2% vs 39.8%).

DISCUSSION

The study found that combining IASTM with standard exercise yielded greater improvements in pain, knee mobility, and function compared with standard exercise alone in individuals with knee OA. This study's findings place IASTM in a clinically valuable position for use alongside regular exercise-based rehabilitation and support a growing IASTM literature base in the musculoskeletal and knee OA populations.

As it relates to pain, the VAS reduction was also significantly greater in Group A than in Group B, in line with the findings of Nazari et al.[16], which found that there was greater pain reduction in the combination group (ISTM + other active treatments) than in the active treatment group alone in upper-body, lower-body, and spinal conditions. The 2025 systematic review and meta-analysis by Tang et al.[17] also showed that combined interventions resulted in greater reductions in pain intensity and pain-pressure threshold than other interventions alone, supporting an additional pain-relieving effect of adding IASTM to other rehabilitation programs. In keeping with this, Gercek et al. [18], in a double-blind randomized controlled trial, showed that even a single IASTM session produced an immediate drop in pain and joint position error among patients with chronic neck pain suggesting that mechanically desensitizing soft tissue can ease pain acutely after just one treatment, a mechanism that could plausibly contribute to the cumulative pain reduction seen in Group A over the six weeks of the present trial. This lines up with Cigdem-Karacay et al.[19], whose randomized, blinded, sham-controlled study found IASTM outperformed a sham treatment for short-term pain relief in myofascial pain syndrome, weighing against a purely placebo-driven explanation. The size of the range-of-motion gain seen in Group A tracks with what Wadhwa et al.[20] reported in their scoping review, namely that IASTM restores function in tight fascia and periarticular soft tissue and thereby boosts joint mobility, matching the sizable between-group gap in flexion ROM and extension deficit found here. The findings also match Gandotra et al.[6], whose work on the iliotibial band and surrounding knee musculature showed improved clinical performance with IASTM, and Karimi Soloklo et al.,[7], whose work found improved gait kinetics after IASTM applied around the knee. The pattern also fits the meta-analysis by Tang et al.,[9], whose meta-analysis found that IASTM produces a meaningful improvement in range of motion across a range of joints and musculoskeletal conditions compared with exercise or no treatment. Collectively, this body of recent evidence points to soft tissue-directed intervention offering a mobility benefit beyond what exercise alone provides.

Group A's WOMAC gains also outpaced Group B's, matching Karimi Soloklo et al.,[8], who found larger gains in lower-limb muscle strength and sit-to-stand ability with IASTM in a comparable knee OA population, and Jafarsalehi et al.,[10], whose trial found greater gains in functional

measures and quality of life when IASTM was added to knee OA treatment.

Additional support comes from the systematic review by Tsokanos et al.[21], which concluded that both manual and instrument-based soft tissue techniques yield clinically meaningful gains in function and disability when added to exercise therapy for knee osteoarthritis. In a similar vein, Mutiara et al.[15] found that manual traction-based soft tissue work significantly improved WOMAC scores in knee OA patients, broadly matching the superior WOMAC gains seen in the IASTM-augmented group here. Taken together, these findings suggest soft tissue-directed adjuncts whether manual or instrument-assisted offer a consistent, incremental functional benefit beyond exercise alone.

A few physiological mechanisms may account for IASTM's superior results in this trial. Brandl et al.[22] used bioimpedance analysis to show that IASTM triggers measurable, immediate shifts in the fluid dynamics and hydration of myofascial tissue, alongside a rise in local tissue temperature consistent with better extensibility and less mechanical restriction. Nazari et al.[16] similarly proposed that IASTM's clinical value stems from a mix of local mechanical effects on fascia and connective tissue plus neurophysiological pain modulation. In a randomized controlled trial that paired targeted IASTM with neuromuscular exercise, Mylonas et al.[23] found that this combination led to significant gains in posture-related dysfunction and pain beyond exercise alone, a synergistic pattern similar to what this trial observed at the knee. These mechanistic and clinical results provide a compelling case for the incremental clinical value of treating periarticular soft tissue restriction in combination with therapeutic exercise rehabilitation in knee OA.

Several methodological details support the confidence in these results. The use of VAS for pain assessment is consistent with the previously mentioned study by Goudman et al.[13], in which the VAS was shown to agree well with the numerical rating scale for measuring pain intensity in adults with chronic pain. The report by Jayubo et al.14 also suggested that universal goniometer-based ROM measurement is reliable; this approach was used in the current study to measure knee flexion and extension. In this trial, which examined the reliability of clinicians' use of IASTM forces during simulated treatment, Duffy et al.[24] found acceptable reliability among trained clinicians, supporting the standardized IASTM approach performed by a single qualified physical therapist.

The strengths of this trial are the randomized parallel-group design, concealed allocation, blinded outcome assessment, and well-defined, repeatable IASTM and exercise protocols. A few limitations are worth flagging too. Because the design was single-blind, treating therapists knew each participant's group assignment, which may have introduced a degree of performance bias. Six weeks of follow-up, though standard for rehabilitation research, is too short to say anything about how long the benefits last, a shortfall also noted by Moore et al.[25] in their evidence-to-practice review, which pointed out a broader shortage of long-term follow-up data across the IASTM literature. Running the trial across two centers also helped feasibility, but it limits how far these results generalize to other clinical settings, patient groups, and age ranges.

CONCLUSION

Overall, adding IASTM to therapeutic exercise led to significantly greater reductions in pain and significantly larger gains in knee range of motion and function than exercise alone over six weeks of treatment for knee osteoarthritis. These results agree with, and add to, the IASTM literature published from 2020 to 2025, pointing to IASTM as a clinically useful and safe addition that fits readily into conventional exercise-based rehabilitation for knee OA. Clinicians treating knee osteoarthritis may want to consider pairing IASTM with therapeutic exercise, since it appears to deliver more pain relief and better restoration of joint mobility and function than exercise on its own. Further research with longer follow-up, recruitment across multiple centers, and blinded treatment providers would help confirm how durable and generalizable these findings are.

Author Contributions

FK: Conception and Design, Data Collection, and Supervision.
MA: Study Design, Data Analysis, and Interpretation**.**
KT: Literature Review, Manuscript Writing, and Critical Review.
RI: Data Collection, Manuscript Editing, and Data Interpretation.
ZM: Manuscript Preparation, Critical Review, and Final Approval.

Funding Source

No sources.

Conflict of Interest

The authors declared no conflict of interest.

Data Availability
Data will be available on request.

Ethical Approval
The protocol was approved by the Institutional Ethical Review Committee at the Foundation of Medical Research and Lab (FMRL-IRB/2023/026).

REFERENCES

1. Chen J, Chen X, Wang T, Li M, Dai H, Shang S, Cheng L, Tang Z, Liu S, Huang C. Global burden of knee osteoarthritis from 1990 to 2021: Trends, inequalities, and projections to 2035. PLoS One. 2025 Jun 11;20(6):e0320115. https://doi.org/10.1371/journal.pone.0320115
2. Giorgino R, Albano D, Fusco S, Peretti GM, Mangiavini L, Messina C. Knee osteoarthritis: epidemiology, pathogenesis, and mesenchymal stem cells: what else is new? An update. International Journal of Molecular Sciences. 2023 Mar 29;24(7):6405. doi: https://doi.org/10.3390/ijms24076405
3. Holden MA, Metcalf B, Lawford BJ, Hinman RS, Boyd M, Button K, Collins NJ, Cottrell E, Henrotin Y, Larsen JB, Master H. Recommendations for the delivery of therapeutic exercise for people with knee and/or hip osteoarthritis. An international consensus study from the OARSI Rehabilitation Discussion Group. Osteoarthritis and Cartilage. 2023 Mar 1;31(3):386-96. doi: https://doi.org/10.1016/j.joca.2022.10.009
4. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, Callahan L, Copenhaver C, Dodge C, Felson D, Gellar K. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis & Rheumatology. 2020 Feb;72(2):220-33. doi: https://doi.org/10.1002/art.41142
5. Jafarsalehi B, Boozari S, Torkaman G. Effects of instrument-assisted soft tissue mobilization combined with exercise on knee joint muscle activity during walking in patients with moderate knee osteoarthritis: A randomized controlled clinical trial. Scientific Reports. 2025 Dec 29;15(1):44743. doi: https://doi.org/10.1038/s41598-025-28418-3
6. Gandotra JB, Ganjave P, Shroff RA. Effect of instrument-assisted soft-tissue mobilisation of iliotibial band with conventional exercise therapy versus conventional exercise therapy alone on clinical performance in patients with knee osteoarthritis: a prospective interventional study. Journal of Clinical and Diagnostic Research. 2025. doi: https://doi.org/10.7860/JCDR/2025/73706.20855
7. Karimi Soloklo Z, Boozari S, Kahrizi S. Effects of a two-week instrument-assisted soft tissue mobilization and exercise therapy versus sham and exercise on gait kinetics in moderate knee osteoarthritis: a randomized controlled trial. Journal of Manual & Manipulative Therapy. 2026 Jan 2;34(1):16-27. doi: https://doi.org/10.1080/10669817.2025.2481594
8. Karimi SZ, Boozari S, Kahrizi S. The effect of short-term instrument-assisted soft tissue mobilization on lower limb muscle strength and sit-to-stand kinetics in patients with knee osteoarthritis: a randomized controlled trial. Middle East Journal of Rehabilitation and Health Studies. 2025;12(3). doi: https://doi.org/10.5812/mejrh-159368
9. Tang S, Sheng L, Xia J, Xu B, Jin P. The effectiveness of instrument-assisted soft tissue mobilization on range of motion: a meta-analysis. BMC Musculoskeletal Disorders. 2024 Apr 23;25(1):319. doi: https://doi.org/10.1186/s12891-024-07452-8
10. Jafarsalehi B, Boozari S, Torkaman G. Exploring the impact of instrument-assisted soft tissue mobilization on functional measures and quality of life in knee osteoarthritis patients: A randomized controlled trial. Middle East Journal of Rehabilitation and Health Studies. 2025;12(1). doi: https://doi.org/10.5812/mejrh-152345
11. Huang K, Cai H. The interplay between osteoarthritis and osteoporosis: Mechanisms, implications, and treatment considerations — A narrative review. Experimental Gerontology. 2024 Nov 1;197:112614. doi: https://doi.org/10.1016/j.exger.2024.112614
12. Liu Y, Wang Y. A comparative study of the efficacy of instrument-assisted soft tissue mobilization and massage techniques in patients with patellofemoral joint pain. Frontiers in Medicine. 2023 Nov 13;10:1305733. doi: https://doi.org/10.3389/fmed.2023.1305733
13. Goudman L, Pilitsis JG, Billet B, De Vos R, Hanssens K, Billot M, Roulaud M, Rigoard P, Moens M. The level of agreement between the numerical rating scale and visual analogue scale for assessing pain intensity in adults with chronic pain. Anaesthesia. 2024 Feb;79(2):128-38. doi: https://doi.org/10.1111/anae.16151
14. Jayubo DT, Demegillo P, Ramas AJ, Apolonio NE, Makabenta K, Aznar ME, Legion DM, Manalo EJ, Canono ML, Rosales J, Sudario GK. Reliability of range of motion assessment using universal goniometer and video goniometer application: a comparative study. ASEAN Journal of Rehabilitation Medicine. 2025;35(2):67-75.
15. Mutiara R, Ardinata D, Fachrina R, Vytra S, Lumongga F. The effect of physiotherapy manual traction techniques on the Western Ontario and McMaster University OA Index (WOMAC) in knee osteoarthritis patients. Jurnal Keperawatan dan Fisioterapi (JKF). 2025 Aug 12;7(2):158-63. doi: https://doi.org/10.35451/jkf.v7i2.2466
16. Nazari G, Bobos P, Lu S, Reischl S, Sharma S, Le CY, Vader K, Held N, MacDermid JC. Effectiveness of instrument-assisted soft tissue mobilization for the management of upper body, lower body, and spinal conditions. An updated systematic review with meta-analyses. Disability and Rehabilitation. 2023 May 8;45(10):1608-18. doi: https://doi.org/10.1080/09638288.2022.2070288
17. Tang S, Sheng L, Wei X, Liang M, Xia J, Chen J. The effectiveness of instrument-assisted soft tissue mobilization on pain and function in patients with musculoskeletal disorders: a systematic review and meta-analysis. BMC Musculoskeletal Disorders. 2025 Mar 14;26(1):257. doi: https://doi.org/10.1186/s12891-025-08492-4
18. Gercek H, Unuvar BS, Umit Yemisci O, Aytar A. Acute effects of instrument assisted soft tissue mobilization technique on pain and joint position error in individuals with chronic neck pain: a double-blind, randomized controlled trial. Somatosensory & Motor Research. 2023 Jan 2;40(1):25-32. doi: https://doi.org/10.1080/08990220.2022.2157388
19. Cigdem-Karacay B, Sahbaz T, Gumusay D, Kodak MI, Tuncay F, Ozudogru A, Kocak FA. Short-term results of instrument-assisted soft tissue mobilization in female patients with myofascial pain syndrome: A randomized blinded sham-controlled study. Turkish Journal of Physical Medicine and Rehabilitation. 2025 Nov 10;71(4):496. doi: https://doi.org/10.5606/tftrd.2025.15466
20. Wadhwa K, Gupta R, Singh S, Kumar M, Singh AK. Impact of instrument-assisted soft-tissue mobilization (IASTM) on the functional status of myofascial structures: scoping literature review. Journal of Iranian Medical Council. 2025 Jun 1;8(3):437-47. doi: https://doi.org/10.18502/jimc.v8i3.18788

21. Tsokanos A, Livieratou E, Billis E, Tsekoura M, Tatsios P, Tsepis E, Fousekis K. The efficacy of manual therapy in patients with knee osteoarthritis: a systematic review. Medicina. 2021 Jul 7;57(7):696. doi: https://doi.org/10.3390/medicina57070696
22. Brandl A, Egner C, Schwarze M, Reer R, Schmidt T, Schleip R. Imediate effects of instrument-assisted soft tissue mobilization on hydration content in lumbar myofascial tissues: a quasi-experiment. Journal of Clinical Medicine. 2023 Jan 28;12(3):1009. doi: https://doi.org/10.3390/jcm12031009
23. Mylonas K, Angelopoulos P, Billis E, Tsepis E, Fousekis K. Combining targeted instrument-assisted soft tissue mobilization applications and neuromuscular exercises can correct forward head posture and improve the functionality of patients with mechanical neck pain: a randomized control study. BMC Musculoskeletal Disorders. 2021 Feb 21;22(1):212. doi: https://doi.org/10.1186/s12891-021-04080-4
24. Duffy S, Martonick N, Reeves A, Cheatham SW, McGowan C, Baker RT. Clinician reliability of one-handed instrument-assisted soft tissue mobilization forces during a simulated treatment. Journal of Sport Rehabilitation. 2022 Feb 2;31(4):505-10. doi: https://doi.org/10.1123/jsr.2021-0216
25. Moore MG, Daley AR, Winkelmann Z. An evidence-to-practice review on the efficacy of instrument-assisted soft tissue mobilization. Clinical Practice in Athletic Training. 2020;3(3):2. doi: https://doi.org/10.31622/2020.0003.3.2

Original Article

Allied Medical Research Journal
Vol:4 No.3: 31-37
© The Author (s) 2026

Frequency of Cervicogenic Headaches symptoms among Bankers of Islamabad: A Cross-Sectional Study

Khaleel Ur Rehman[1], Eeman Malik[2], Mahpara Ali[2]

[1]*Sheffield Hallam University, England, United Kingdom*
[2]*Zulfiqar Ali Bhutto Medical University, Islamabad, Pakistan*

ABSTRACT

Background: Cervicogenic headache (CGH) is a secondary headache associated with disorders of the cervical spine. Poor and sustained posture, particularly during prolonged desk work and use of electronic devices, may contribute to musculoskeletal pain, reduced cervical mobility, sleep disturbances, and fatigue. Bankers are frequently exposed to these occupational risk factors. This study aimed to determine the frequency and severity of CGH symptoms among bankers working in Islamabad.

Methods: A cross-sectional descriptive study was conducted among 61 bankers aged 25–60 years working in banks in Islamabad. Participants were recruited using convenience sampling. Data were collected from July 2025 to January 2026 using the Visual Analog Scale (VAS) to assess pain intensity and the Wiley-Blackwell Cervicogenic Headache Diagnostic Criteria Questionnaire. The questionnaire was used to assess factors associated with CGH and was not intended as a substitute for a comprehensive clinical diagnosis. Data were analyzed using frequencies and percentages in SPSS version 23.

Results: CGH was identified in 20 of 61 participants, indicating a prevalence of 32.7%. Among those meeting the criteria for CGH, 10% reported mild pain, 65% reported moderate pain, and 25% reported severe pain. Moderate pain was therefore the most frequently reported intensity.

Conclusion: A notable proportion of bankers reported symptoms consistent with CGH, with moderate pain being the predominant severity. These findings highlight the importance of workplace ergonomics, posture awareness, and preventive strategies to reduce occupational risk factors associated with CGH among desk-based professionals.

Keywords: *Bankers; Cervicogenic Headache; Ergonomics; Headache Disorders; Posture.*

Received: May 07th, 2026**; Revised:** June 19th, 2026**; Accepted:** June 26th, 2026

Corresponding Email: khaleelrehman797@gmail.com

DOI: https://doi.org/10.59564/amrj/04.03/004

INTRODUCTION

Headaches have been experienced in about 47% of people worldwide[1]. Among them, 2% to 3% of people worldwide experience cervicogenic headaches (CGH)[2]. Clinicians from a variety of specializations and professions have been describing CGH for a long time[3]. According to the literature, there has been a long-standing debate about whether headaches are related to cervical symptoms[4]. In comparison, some experts consider CGH as a separate condition while others contend that it is actually a part of other headache disorders like tension-type headache and migraine[5-6].

A CGH originates from the upper cervical vertebra due to the compression of the sensory nerves supplying the scalp[7]. A typical pattern of pain is stiffness in the muscles surrounding the neck and upper cervical vertebrae, followed by a rise in headache in the posterior scalp, and later progressing to the anterior, including diplopia, dizziness, dysphagia, drop attacks, and dysarthria[1,8]. The fascia, disc, and soft tissue surrounding the neck are all sensitive components that might cause the CGH[9,10]. Cluster headache and migraine are frequently misdiagnosed in patients with CGH[11]. The head

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

pain experienced by many patients with cervical spondylosis is also cervicogenic[12]. Certain traits help in distinguishing CGH from other headaches, including side-locked pain, which gets worse with head movement[13,14].

Multiple international groups, including the International Headache Society and the Cervicogenic Headache International Study Group, proposed the diagnostic criteria for CGH[15]. Some of the criteria include radiographic evidence of abnormalities or diagnostic nerve blocks such as ICHD3[16]. Physiotherapy administration also significantly impacts the treatment of CGH[17]. Effective treatment options for the cervical spine include manual therapy, targeted therapeutic exercises, spinal traction, and range-of-motion exercises[18]. Cervical pain has multiple causes, but office employees who sit for more than 6 hours are more likely to experience it due to prolonged sitting with improper posture, reduced postural control, and increased physical and mental strain from work[19].

Due to the prolonged work hours and extensive computer use in the banking industry, neck pain is commonly experienced[20]. To promote awareness, provide interventions, and enhance educational programs, it is crucial to understand the dimensions and characteristics of CGH among computer users [21]. The issue of CGH is addressed frequently among computer users, but not specifically among bankers [20]. The current research was carried out to fill the gap in the literature on CGH among bankers in Islamabad, as very little literature exists on this musculoskeletal problem in our country. The purpose of the research was to identify the frequency of CGH among bankers. The study will add valuable information to the understanding, prevention, and management of CGH among bankers by determining its prevalence in this population, thereby enhancing bankers' quality of life.

METHODOLOGY

The research was a descriptive cross-sectional study conducted over seven months, from July 2025 to January 2026, across various banks in Islamabad. The target population was computer-using employees. The sample comprised 61 bankers, calculated using OpenEpi software 3.0. For prevalence studies, assuming an expected prevalence of 42.4% based on a previous study among bankers, a 95% confidence level, and a 5% margin of error, sampled through a non-probability convenience sampling method. The two instruments used were the Wiley-Blackwell Cervicogenic Headache: Diagnostic Criteria questionnaire and the Visual Analog Scale (VAS) for pain assessment. The study included bank employees of both sexes, with 91.8% male, aged 25-60 years, working more than 8 hours a day, including more than 5 hours at a computer. Participants were screened for the characteristics suggestive of CGH using the Wiley-Blackwell Cervicogenic Headache Diagnostic Criteria Questionnaire. Those who had experienced a history of traumatic injuries, neck or upper limb surgery, medical conditions (neuropathy in diabetic patients or sensory loss), and chronic diseases (rheumatoid arthritis or osteoarthritis) were excluded. The diagnostic questionnaire was used to screen the participants.

Ethical approval was obtained from the ethics committee of the Foundation of Medical and Research and Laboratories (Protocol no. FMRL-IRB/2024/028). After ethical approval was obtained, bankers from different workplaces were approached during working hours. The purpose of the study had been explained previously. Written informed consent was also obtained before data collection. Participants meeting the inclusion criteria were required to complete the questionnaire.

Data had been collected between July 2025 and January 2026. The Statistical Package for the Social Sciences (SPSS) version 23.0 was used to analyze the collected data. Frequency graphs were created for qualitative variables, and descriptive statistics were computed for quantitative variables (VAS scores>3).

RESULTS

The findings showed that 61 participants were part of the sample. The age group (25 to 60) outlined in the research was further classified as follows: (25 – 32), (33-39), (40 – 46), (47- 53), and (54-60). Of the total sample size of 61 participants, the largest share, i.e., 37.7%, fell in the age group (25-32), and the smallest share, i.e., 4.92%, fell in the age group (54-60). The total number of participants by gender is shown in Figure 1. The total numbers of male and female patients were 91.8 and 8.2, respectively, indicating a higher ratio of male patients experiencing headache.

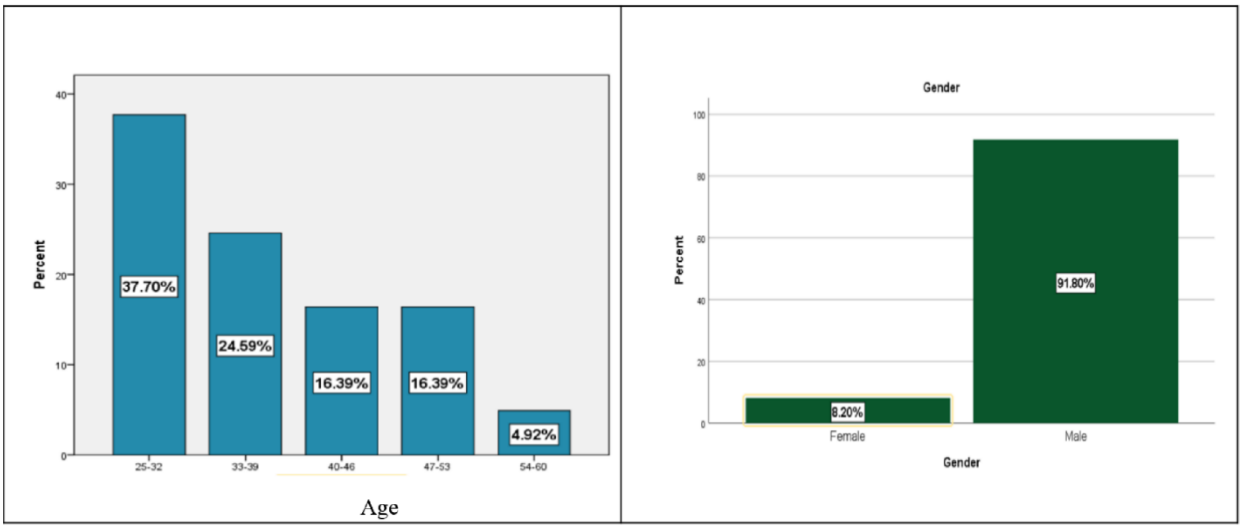

Figure 1: Demographic Variables of Participants

Table 1 presents the years of work experience and the total daily working hours of bankers in Islamabad. The work experience of bankers was assessed in four categories, ranging from 1 year to 4 years. Of 61 individuals, most bankers [n=26, 42.6%] had 1 year of work experience, whereas at least [n=8, 13.1%] had 4 years of experience. Working hours were also assessed in three categories, ranging from 8 to 10 hours, and most bankers [n=38, 62.3%] have 9 hours of daily work; at least [n=4, 6.6%] work 10 hours.

Table 1: Working Experience and Working Hours of Bankers [N= 61]

Variable	Category	Frequency	Percent
Working Years of participants	1 year	26	42.6
	2 years	17	27.9
	3 years	10	16.4
	4 years	8	13.1
Daily working hours	8 hours	19	31.1
	9 hours	38	62.3
	10 hours	4	6.6

Figure 2a shows the distribution of pain severity among subjects who met criteria suggestive of CGH. Subjects reporting moderate pain intensity (65%), followed by severe pain (25%) and mild pain (10%).

Figure 2b shows that 20 of the 61 participants have fulfilled the criteria suggestive of CGH, yielding a prevalence of 32.7% (95% CI: 21.0%–44.6%).

Table 2 below shows the frequency distribution of the clinical characteristics of cervicogenic headache.	
Factors	**F (%)**
1. One sided headache without side-shift	22 (36.1%)
2a. Symptoms showing the involvement of neck: Pain initiated by movement of neck/sustained awkward posture	21 (34.4 %)
2b. Involvement of neck indicating symptoms: External pressure of the posterior neck or occipital region	7 (11.5%)
2c. Symptoms and signs of neck involvement: Ipsilateral neck, shoulder, and pain in arm	14 (23%)
2d. Symptoms and signs of neck involvement: Decreased range of Motion	13 (21.3 %)
3. Pain episodes of varying duration or fluctuating continuous pain	26 (42.6%)
4. Moderate, non-excruciating pain, usually of a non-throbbing nature	28 (45.9%)
5. Pain starting in the neck, spreading to oculo-fronto-temporal areas	20 (32.8%)
6a. Various attack-related events: autonomic symptoms and signs, Nausea, vomiting, dizziness	11 (18%)
6b. Various attack-related events: autonomic symptoms and signs, Ipsilateral edema	3 (4.9%)
6c. Various attack-related events: autonomic symptoms and signs, Flushing in the peri-ocular area	22 (36.1%)
6d. Various attack-related events: autonomic symptoms and signs, Photophobia, phonophobia.	9 (14.8%)
6e. Various attack-related events: autonomic symptoms and signs, Blurred vision in the Ipsilateral eye	22 (36.1%)

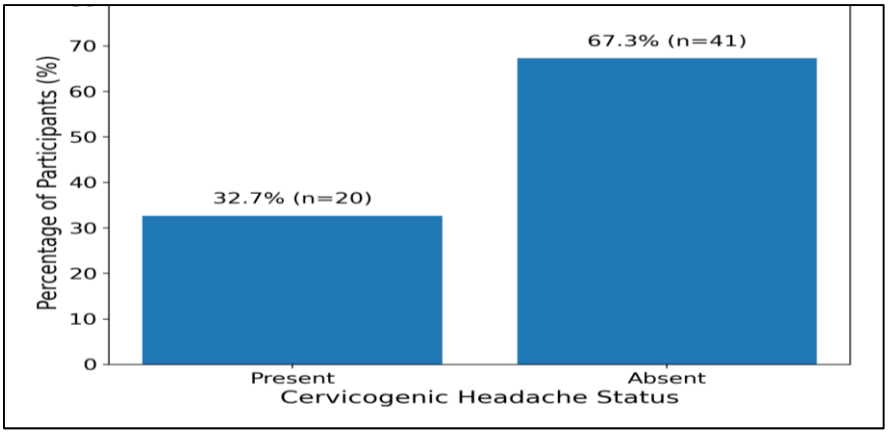

Figure 2a: Severity of Pain among Participants with Cervicogenic Headache

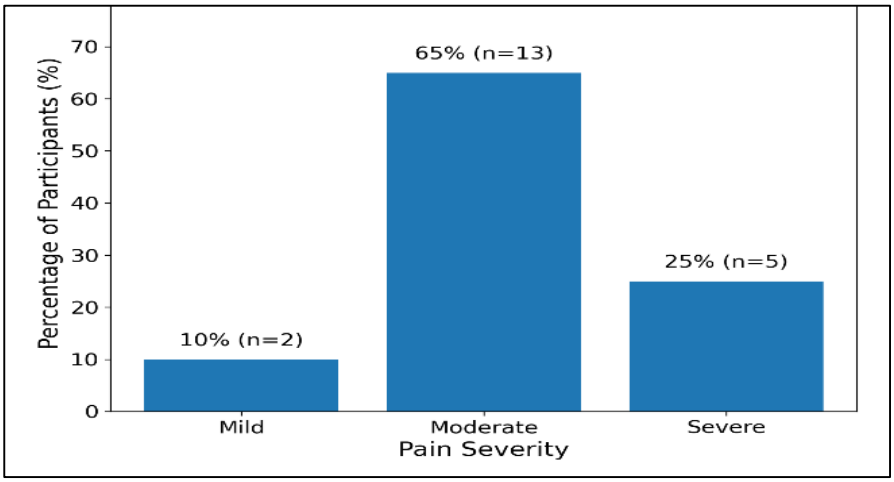

Figure 2b: Prevalence of Cervicogenic Headache among Study Participants

DISCUSSION

The results of the current study showed that CGH is prevalent among bankers working in banks in Islamabad. Current results obtained using the Visual Analog scale and the Wiley-Blackwell Cervicogenic Headache: Diagnostic Criteria indicate that CGH is present among bankers. There is evidence found regarding cervical pain among bankers due to abnormal posture, occupational hazards, and job factors that include long working hours, hectic routine, workload, age, but no such study has been found in Islamabad[22].

A survey was conducted among computer users at the University of Sargodha to determine the prevalence of CGH and neck pain. Using a self-developed questionnaire distributed via email, 119 responses were collected. The findings showed that 73.1% of participants experienced pain on most days, with 58% reporting pain in the head and neck region. Additionally, 57.1% used chairs and sofas with back support, while 108 individuals attributed their pain to incorrect posture. These results indicate a growing prevalence of head and neck pain among computer users, likely due to prolonged use of laptops and computers in incorrect postures[23]. A cross-sectional study was conducted to determine the prevalence of neck pain among bankers. A self-designed questionnaire was administered to 416 participants, both males and females aged 18-55 years. The bankers' work duration ranged from 8 to 12 hours per day. The study findings indicated a prevalence of 42.4% for neck pain among bankers, with 26.99% in males and 15.42% in females, reflecting the male-dominant data collection[24]. These findings highlight the importance of addressing factors such as prolonged working hours and promoting ergonomic practices in the banking industry to reduce neck pain among bankers and improve their overall well-being. Further research is needed to explore additional contributing factors and evaluate the effectiveness of preventive measures[24].

Our study reported a prevalence of CGH of 32.7%, lower than that reported in the referenced study, possibly due to differences in the study population and sampling methods.

To determine the frequency and factors associated with headaches among individuals in various occupations, a cross-sectional study was conducted in four cities: Karachi, Rawalpindi, Rahim Yar Khan, and Muzaffarabad. The study included 250 individuals aged 18-60 years from various professions, including doctors, engineers, businesspeople, executives, drivers, teachers, armed officers, laborers, and household servants[25]. In another study conducted in 2013 by Huber et al.[26] titled "Reinvestigation of the Dysfunction in Neck and Shoulder Girdle Muscles as the Reason for Cervicogenic Headache among Office Workers," They conducted a comprehensive investigation to examine the association of CGH in office workers. The findings highlighted significant links between these muscular dysfunctions and the occurrence of CGH among office workers.

The research has some limitations, such as a small sample size, which can increase the risk of error; the lack of an imaging diagnostic test; and the short duration of the research, which may limit the validity of the prevalence results. Also, the number of individuals approached, declined, or excluded had not been recorded systematically, limiting the response rate. Also, the predominance of male participants (91.8%) may limit the generalizability of the findings to female bankers. Irrespective of these limitations, the research has meaningful information and can be useful to patients and physiotherapists in CGH management.

Future studies are to be done on a multi-centric, longitudinal, and cross-sectional basis with more contributing factors like depression, psychological problems, occupational hazards, and other demographic variables. In addition, CGH in female bankers who resume work after giving birth should also be investigated in the future.

CONCLUSION

The results showed that the most common related symptoms were pain episodes of varying lengths, moderate, non-throbbing pain, and neck pain extending to the oculo-fronto-temporal area, indicating that occupational posture and neck strain may be associated with the CGH.

Author Contributions

KUR: Conception and Design, Data Collection, and Supervision.
EM: Study Design, Data Analysis, and Interpretation.
MA: Literature Review, Manuscript Writing, and Critical Review.

Funding Source

No sources.

Conflict of Interest
The authors declared no conflict of interest.

Data Availability
Data will be available on request.

Ethical Approval
It was obtained from the ethical committee of Foundation of Medical and Research and laboratories (Protocol no. FMRL-IRB/2024/028).

REFERENCES

1. Al Khalili Y, Ly NK, Murphy PB. Cervicogenic Headache. [Updated 2022 Oct 3]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK507862/
2. de Wijer A, de Leeuw JRJ, Steenks MH, Bosman F. Temporomandibular and cervical spine disorders: self-reported signs and symptoms. Spine. 1996;21(14):1638-46. doi: https://doi.org/10.1097/00007632-199607150-00005
3. Jin X, Zhang Q, Lin J, Sun Y, Dong Z, Jin W, et al. Minimally Invasive Interventions for Cervicogenic Headache: A Systematic Review of Targeted Anatomical Approaches and Clinical Outcomes. J Pain Res. 2026 Jan 8;19:549239. doi: https://doi.org/10.2147/JPR.S549239
4. Vicente BN, Oliveira R, Martins IP, Gil-Gouveia R. Cranial Autonomic Symptoms and Neck Pain in Differential Diagnosis of Migraine. Diagnostics (Basel). 2023 Feb 5;13(4):590. doi: https://doi.org/10.3390/diagnostics13040590
5. Onan D, Younis S, Wellsgatnik WD, Farham F, Andruškevičius S, Abashidze A, et al. Debate: differences and similarities between tension-type headache and migraine. J Headache Pain. 2023 Jul 21;24(1):92. doi: https://doi.org/10.1186/s10194-023-01614-0
6. Pescador Ruschel MA, De Jesus O. Migraine Headache. [Updated 2024 Jul 5]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560787/
7. Aabrooi S, Shafique S, Javed A, Fatima A, Khan O, Riaz S. Frequency of cervicogenic headache in students due to usage of smart devices-cross sectional survey. Pak J Rehabil. 2022;11(2):36-41. doi: https://doi.org/10.36283/pjr.zu.11.2/007
8. Graff-Radford SB. Facial pain, cervical pain, and headache. Continuum (Minneap Minn). 2012;18(4):869-82. doi: https://doi.org/10.1212/01.con.0000418648.54902.42
9. Afzal F, Atif MM, Kumar A, Rasul A, Islam A, Nadeem I. Cervicogenic headache and neck pain among computer users. Rawal Med J. 2020;45(2):370-2. Available from: https://api.semanticscholar.org/CorpusID:219434371
10. Šedý J, Rocabado M, Olate LE, Vlna M, Žižka R, et al. Neural Basis of Etiopathogenesis and Treatment of Cervicogenic Orofacial Pain. Medicina (Kaunas). 2022 Sep 21;58(10):1324. doi: https://doi.org/10.3390/medicina58101324
11. Jain H, Pachpute S. Prevalence of Cervicogenic Headache in Computer Users. AIJR Abstr. 2022:56. Available from: https://api.semanticscholar.org/CorpusID:251515250
12. Yabuki S, Takatsuki K, Otani K, Nikaido T, Watanabe K, Kato K, et al. Headache in Patients with Cervical Spondylotic Myelopathy. Pain Res Manag. 2020 Sep 28;2020:8856088. doi: https://doi.org/10.1155/2020/8856088
13. Lefel N, van Suijlekom H, Cohen SPC, Kallewaard JW, Van Zundert J, et al. Cervicogenic headache and occipital neuralgia. Pain Pract. 2025 Jan;25(1):e13405. doi: https://doi.org/10.1111/papr.13405
14. Li J, Wu B, Wang X, Zhao L, Cui J, Liu J, Guo K, Zhang, et al. Clinical features, risk factors, and a nomogram for predicting refractory cervicogenic headache: a retrospective multivariate analysis. Front Neurol. 2025 Mar 24;16:1531180. doi: https://doi.org/10.3389/fneur.2025.1531180
15. McGough S, Fernandes L, Idrovo L. Peripheral nerve blocks for primary and secondary headache disorders: review of current evidence and a practical approach. J Oral Facial Pain Headache. 2026 Jan;40(1):25-31. doi: https://doi.org/10.22514/jofph.2026.002
16. Li D, Jin X, Li J, Li M, Jin S, Jin W. Combined ultrasound-guided C2 DRG pulsed radiofrequency and suboccipital myofascial plane block for cervicogenic headache: a retrospective study. Front Pain Res (Lausanne). 2026 Jan 26;7:1725798. doi: https://doi.org/10.3389/fpain.2026.1725798
17. Jung A, Carvalho GF, Szikszay TM, Pawlowsky V, Gabler T, Luedtke K. Physical Therapist Interventions to Reduce Headache Intensity, Frequency, and Duration in Patients With Cervicogenic Headache: A Systematic Review and Network Meta-Analysis. Phys Ther. 2024 Feb 1;104(2):pzad154. doi: https://doi.org/10.1093/ptj/pzad154
18. Al Khalili Y, Ly NK, Murphy PB. Cervicogenic Headache. [Updated 2022 Oct 3]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK507862/
19. Umer M, Motuma A, Ahmed N, Letta S. Prevalence and Associated Factors of Work-Related Spinal Pain Among Bank Workers in Harar, Eastern Ethiopia: A Cross-Sectional Study. SAGE Open Nurs. 2025 Aug 21;11:23779608251371100. doi: https://doi.org/10.1177/23779608251371100
20. Marzban H, Rezaei E, Shahmahmoudi F, Zangiabadi Z, Sahebi A, Makki F, et al. Musculoskeletal disorders among bank workers: a systematic review and meta-analysis. BMC Musculoskelet Disord. 2024 Nov 25;25(1):951. doi: https://doi.org/10.1186/s12891-024-08077-7
21. Verville L, Dc PC, Grondin D, Mior S, Moodley K, Kay R, et al. Using technology-based educational interventions to improve knowledge about clinical practice guidelines. J Chiropr Educ. 2021 Mar 1;35(1):149-57. doi: https://doi.org/10.7899/jce-19-17
22. Shabbir M, Rashid S, Umar B, Ahmad A, Ehsan S. Frequency of neck and shoulder pain and use of adjustable computer workstation among bankers. Pak J Med Sci. 2016 Mar-Apr;32(2):423-6. doi: https://doi.org/10.12669/pjms.322.9067
23. Boota M, Shaban S, Munawar A, Ahmed A. Prevalence of neck pain in bankers of Lahore. Malays J Med Res. 2018;2(1):30-6. Available from: https://ejournal.lucp.net/index.php/mjmr/article/view/243
24. Herekar AA, Ahmad A, Uqaili UL, Ahmed B, Effendi J, Alvi SZ, et al. Primary headache disorders in the adult general population of Pakistan - a cross sectional nationwide prevalence survey. J Headache Pain. 2017 Dec;18(1):28. doi: https://doi.org/10.1186/s10194-017-0734-1

25. Huber J, Lisiński P, Polowczyk A. Reinvestigation of the dysfunction in neck and shoulder girdle muscles as the reason of cervicogenic headache among office workers. Disabil Rehabil. 2013 May;35(10):793-802. doi: https://doi.org/10.3109/09638288.2012.709306

Original Article

The Impact of DBT-Based Mindfulness Training on Cognitive Performance among University Students: A Pre-Post Stroop Test Study

Allied Medical Research Journal
Vol:4 No.3: 38-44
© The Author (s) 2026

Shumaila Atif [1], Sheeba Farhan[2], Rabika Shoukat[1], Sara Jehangir[1]

[1]*College of Clinical Psychology, Faculty of Allied Health Sciences, Ziauddin University, Karachi, Pakistan*
[2]*Assistant Professor, Department of Psychology, FUUAST, Abdul Haq Campus*

ABSTRACT

Background: Mindfulness is increasingly used to support cognitive and emotional functioning in high-stress academic settings. Dialectical Behavior Therapy (DBT) combines mindfulness with distress tolerance, cognitive flexibility, and emotion regulation skills that may strengthen attentional control. However, controlled evidence using objective cognitive measures remains limited in South Asian university populations.

Methods: A one-group pre-test–post-test design was used with 50 undergraduate students (aged 18–25 years) purposively recruited from a Pakistani university. Participants completed the Stroop Color and Word Test (SCWT) before and after a structured six-session DBT-based mindfulness training program covering core mindfulness, distress tolerance, cognitive flexibility, and stress-management skills. Pre- and post-intervention scores on nine SCWT indices were compared using paired-samples t-tests.

Results: Word-reading time fell significantly, from 15.19±5.19 to 12.89±3.93 seconds (t(49)=2.691, p=0.010), as did letter-reading time, from 32.38±10.07 to 26.10±9.33 seconds (t(49)=3.596, p=0.001), and letter-reading errors, from 0.74±0.99 to 0.40±0.78 (t(49)=2.084, p=0.042). Color-naming time, the remaining error scores, and all three interference (conflict) scores showed no significant change (all p>0.05).

Conclusion: A brief DBT-based mindfulness program improved basic processing speed and selective attention but left higher-order executive control and interference resolution largely unchanged. The training is a feasible, low-cost way to strengthen foundational attentional processes in university students. However, longer or more intensive programs may be needed to affect more complex executive functions.

Keywords: *Attention; Executive Function; Dialectical Behavior Therapy; Mindfulness; Self-regulation.*

Received: June 5th, 2026**; Revised:** July 15th, 2026**; Accepted:** July 20th, 2026

Corresponding Email: shumaila.atif@zu.edu.pk

DOI: https://doi.org/10.59564/amrj/04.03/005

INTRODUCTION

University life brings academic workload, performance pressure, and major life changes, all of which draw heavily on students' attention and self-regulation. Mindfulness-based interventions have become common in higher education for exactly this reason, and meta-analytic evidence shows they produce small-to-moderate improvements in student distress, anxiety, and dispositional mindfulness compared with passive controls[1]. Related work suggests that mindfulness training also builds resilience, which in turn softens the psychological impact of academic stress[2].

Mindfulness does more than ease distress; it has also been linked to change in core cognitive processes such as attention and cognitive flexibility. A meta-analysis of 87 studies found that meditation practice most reliably strengthens the brain's executive-control network, with weaker and less consistent effects on alerting and orienting[3], and cognitive flexibility itself is now understood as a trainable capacity supported by fronto-parietal brain networks[4]. Even a single guided mindfulness session has been shown to speed up Stroop task reaction times in people with no meditation experience[5], and comparable single-session studies have found matching gains in conflict-resolution speed alongside changes in frontal neural activity measured through functional near-infrared spectroscopy[6].

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

Dialectical Behavior Therapy (DBT) pair mindfulness with distress tolerance, cognitive flexibility, and emotion regulation skills, and its group-based skills format has moved well beyond its clinical origins into non-clinical, help-seeking groups[7]. Reviews of DBT skills groups report modest but consistent benefit across several stress- and regulation-related outcomes[8], and stand-alone DBT skills protocols have reduced impulsivity-related behavior in non-clinical adult samples[9,10]. The mindfulness module of DBT trains people to notice thoughts and sensations without judgment and to keep returning attention to the present moment, what DBT calls the "Wise Mind." That skill set maps fairly directly onto what the Stroop Color and Word Test (SCWT) demands: overriding an automatic reading response in favor of a slower, less automatic color-naming response, which is essentially a behavioral test of selective attention and inhibitory control.

Even so, the evidence on whether brief mindfulness training changes objective cognitive performance, rather than just self-reported mindfulness or distress, is mixed. Several brief, online, or single-module programs report gains in processing speed and attention with no matching change in higher-order executive control[11-13], and a recent 31-day randomized trial found that an online program improved mindfulness and metacognition but produced no measurable change on any cognitive or affective outcome[14]. A systematic review of mindfulness-based interventions in tertiary students reached a similar conclusion, describing cognitive benefits as inconsistent across trials and generally backed by low-quality evidence[15].

In Pakistan, adapted mindfulness training has already shown good feasibility and acceptability in reducing student stress[16]. However, controlled evidence of its cognitive effects, measured objectively rather than through self-report, remains limited. This study examines whether a brief, structured DBT-based mindfulness training program changes cognitive performance, specifically processing speed, selective attention, and interference control, in university students, using a pre-test–post-test design with the SCWT as the main outcome measure.

METHODOLOGY

Design

This study used a one-group pre-test–post-test design to track changes in cognitive performance following a structured DBT-based mindfulness training program, comparing each participant's own scores before and after training rather than against a separate control group.

Participants

Fifty undergraduate students aged 18–25 were recruited from a Pakistani university through purposive sampling. To be included, students had to be enrolled full-time, fall within the target age range, and have no history of neurological illness or psychotropic medication use. Anyone with a current, diagnosed neurological or psychiatric condition, or prior formal mindfulness or DBT training, was excluded. The final sample (24 male, 26 female) spanned all four undergraduate years and four broad fields of study (Table 1); most participants lived with family (72%) in nuclear households (62%).

Measures

Cognitive performance was measured with the Stroop Color and Word Test (SCWT), a long-established neuropsychological measure of selective attention, processing speed, and inhibitory control. Participants read printed words, read color names printed in mismatched ink, and named the ink color itself while trying to suppress the more automatic reading response[6,14]. This produced nine indices in total: time and error scores for word reading, letter reading, and color naming, plus three interference (conflict) scores capturing the cost of overriding the dominant reading response.

Intervention

The intervention ran across six weekly sessions of about 90 minutes each, preceded by a baseline orientation session. Training moved from core mindfulness skills and "Wise Mind" awareness, through distress-tolerance skills (STOP, TIPP), cognitive flexibility and mindful decision-making exercises, and finally to stress management and relaxation techniques such as progressive muscle relaxation and paced breathing, and concluded with a consolidation and evaluation session. Combining mindfulness with distress-tolerance and emotion-regulation modules in this way is the standard DBT skills-training format, and it has already shown measurable benefit for impulsivity- and stress-related outcomes in non-clinical and student samples[8,9]. Participants kept homework logs between sessions to support daily practice, and each session began with a brief group review of what they had noticed since the last session.

Procedure

The SCWT was administered individually in a quiet, distraction-free room before the first training session and again after the sixth session, by an assessor who was not involved in delivering the sessions. Participants were identified only by code throughout the study, and they were free to withdraw at any point without penalty.

Ethical Considerations

Prior to data collection, the study protocol was approved by the Foundation Research Forum Institutional Review Board (IRB No. 2025/056). Before participating in the study, all participants were informed of its purpose, procedures, and voluntary nature, and then signed informed consent forms.

All data were coded rather than named, and confidentiality was respected during data processing and analysis; data were processed only for research purposes.

Statistical Analysis

The data was analyzed using SPSS. Sample descriptive (means, standard deviations, and frequencies) were used to summarize the sample, and paired-samples t-tests were used to compare the pre- and post-intervention scores across each of the nine indices of the SCWT, with a significance level of $p<0.05$.

RESULTS

Table 1 summaries the sample's demographic profile (N=50). Participants were fairly evenly spread across academic years, close to an even gender split (48% male, 52% female), and represented social sciences, business and management, medical and health sciences, and computer sciences.

Table 1. Demographic Characteristics of Participants (N = 50)

Variable	Category	n	%
Age Group	18–20 years	18	36.0
	21–23 years	22	44.0
	24–25 years	10	20.0
Gender	Male	24	48.0
	Female	26	52.0
Academic Year	1st Year	12	24.0
	2nd Year	15	30.0
	3rd Year	13	26.0
	4th Year	10	20.0
Field of Study	Social Sciences	16	32.0
	Business & Management	12	24.0
	Medical & Health Sciences	11	22.0
	Computer Sciences	11	22.0
Family System	Nuclear	31	62.0
	Joint	19	38.0
Residence Status	Hostel	14	28.0
	Living with Family	36	72.0

Table 2 sets out the paired-samples t-tests comparing pre- and post-intervention SCWT scores. Word-reading time dropped from 15.19±5.19 seconds at pre-test to 12.89±3.93 seconds at post-test ($t(49)=2.691$, $p=0.010$), and letter-reading time dropped from 32.38±10.07 to 26.10±9.33 seconds ($t(49)=3.596$, $p=0.001$). Letter-reading errors

also fell, from 0.74±0.99 to 0.40±0.78 (t(49)=2.084, p=0.042). Color-naming time decreased but not significantly (t(49)=1.435, p=0.158), and neither word-reading errors (p=0.860) nor color-naming errors (p=0.288) changed. None of the three interference (conflict) scores, for words, letters, or color, shifted significantly either (all p>0.05), though the letter-conflict score came close to a significant increase (t(49)=-1.941, p=0.058).

Table 2. Paired-Samples t-Test for Pre- and Post-Intervention Stroop Color and Word Test Scores (N = 50)

Variable	Pre-Test M (SD)	Post-Test M (SD)	t (49)	p
Time – Read Words (s)	15.19 (5.19)	12.89 (3.93)	2.691	0.010*
Time – Read Letters (s)	32.38 (10.07)	26.10 (9.33)	3.596	0.001**
Time – Say Colour (s)	23.14 (9.28)	20.98 (6.51)	1.435	0.158
Errors – Read Words	0.24 (0.52)	0.22 (0.55)	0.178	0.860
Errors – Read Letters	0.74 (0.99)	0.40 (0.78)	2.084	0.042*
Errors – Read Colour	0.96 (1.11)	0.70 (1.34)	1.074	0.288
Conflict – Words	1.04 (0.20)	1.06 (0.24)	-0.444	0.659
Conflict – Letters	1.06 (0.24)	1.16 (0.37)	-1.941	0.058
Conflict – Colour	1.40 (0.50)	1.46 (0.50)	-0.724	0.472

**p<0.05, **p<0.01*

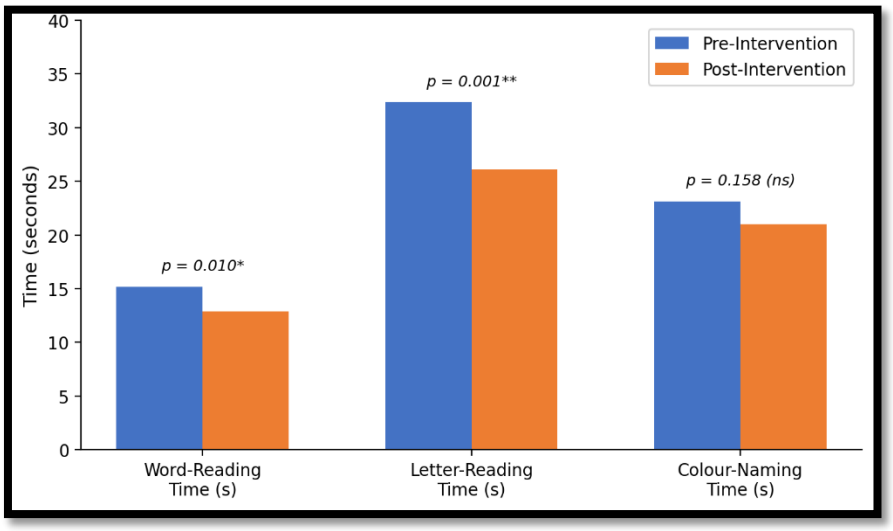

Figure 1 shows this pattern across the three timed indices

Word-reading and letter-reading times both fell after the intervention, while color-naming time barely moved, suggesting the benefit sat mainly in the more automatic reading tasks rather than the harder color-naming task (Figure 1).

Combined, this suggests a moderate improvement in basic processing and selective attention (word- and letter-reading time) and in letter-reading errors but not in any corresponding measures of color-naming, in the more challenging errors, or in interference resolution.

DISCUSSION

This study aimed to determine whether a brief, structured mindfulness training program based on DBT affects objective cognitive function in university students. The first of these hypotheses was supported: training was significantly correlated with enhanced processing speed and selective attention in both word and letter reading, as well as fewer letter-reading errors. The second hypothesis was only partially supported, as the scores for color-naming and the more difficult error scores did not increase or decrease, nor did the three interference scores, indicating that the training

did not significantly impact higher-order executive control.

These results fit a wider, and increasingly mixed, literature on mindfulness and cognition. Brief and digital mindfulness formats have repeatedly helped student mental health and self-reported attention problems[17-21], and a similar short programme for nursing students improved perceived stress and dispositional mindfulness over eight weeks[12]. But direct transfer to standardized cognitive tasks looks less dependable: a recent 31-day randomized trial found gains in mindfulness and metacognition but nothing on any cognitive or affective measure[14], and a systematic review of mindfulness interventions in tertiary students came to a similar conclusion, calling the cognitive evidence inconsistent and generally weak[15]. What we found here, improvement on the simpler, more automatic reading measures but not on colour-naming or conflict resolution, fits that pattern, and may simply reflect how short six sessions is next to the sustained practice usually needed to move higher-order inhibitory control[13].

The gain on basic processing measures makes sense given what the DBT mindfulness module actually trains: present-moment, non-judgmental observation, which is close to the sustained, undistracted attention that speeds up simple word- and letter-reading. Single-session and brief mindfulness protocols have produced similar acute gains in reaction time and conflict-resolution speed on the Stroop task in non-clinical samples[22-23], which supports this explanation. The lack of change on interference scores fits, too; mindfulness-based stress-reduction work has reduced experiential avoidance and impulsivity without necessarily touching more effortful control processes over a short intervention window[24]. A comparable pattern turns up in a randomized trial of mindfulness-based stress reduction, which protected academic persistence without changing how students appraised academic stressors, suggesting mindfulness training can help with specific, narrow outcomes without reshaping everything about how students function[25].

Strengths

Compared with subjective self-report, this study employed an objective, standardized neuropsychological measure (SCWT); provided a manualized DBT intervention with explicit session content; and employed a within-participant design that controlled for differences in cognitive ability among individuals at baseline. The short duration and group nature of the program also make it possible to run it as a routine university counseling service.

Limitations

No control/comparison group was used in this study, reducing the ability to determine cause-and-effect; some of the improvement may be due to practice effects (taking the SCWT twice) rather than the intervention. The findings are not generalizable beyond the one sampled institution, were non-random, and only six sessions were conducted, which may not have been sufficiently long to measure change in higher-order executive functions. Similarly, using only one behavior, without a self-report or a neurophysiological measure to accompany it, provides limited insight into the results.

Recommendations

Future research should employ randomized designs with active or waitlist comparison groups to control for differences due to repeated testing and to distinguish the effect of the intervention itself from that of repeated testing. The longer, more intensive DBT-based mindfulness programs should be studied to determine whether changes in interference control and other higher-order executive functions are possible after sustained practice. Larger samples with more sites across varied disciplines and institutions would improve generalizability, and the SCWT could be combined with self-report mindfulness measures and, if available, neurophysiological measures (e.g., EEG or fNIRS) to provide more information about cognitive change.

CONCLUSION

This study showed that a six-session, short DBT-based mindfulness training program had a significant effect on basic processing speed and selective attention performance, measured by the Stroop Color and Word Test word reading and letter reading times, but had a negligible effect on interference control and color naming times, the more challenging error measures, in university students. Teaching basic attentional skills with DBT might be a viable and cost-effective intervention in academic contexts, whereas it is premature to overestimate its efficacy for higher-order executive functions. It is reasonable to suggest that a short mindfulness intervention, developed using DBT, can be incorporated into counseling and study-skills courses as interventions to support students' cognitive and emotional well-being; however, longer and more controlled trials are required to determine if longer-term mindfulness practice can expand

these benefits to more complex executive-control processes that have a greater impact on academic outcomes.

Author Contributions

SA: Conception and Design, Data Collection, and Supervision.
SF: Study Design, Data Analysis, and Interpretation.
RS: Literature Review, Manuscript Writing, and Critical Review.
SJ: Data Collection, Manuscript Editing, and Final Approval.

Funding Source

No sources.

Conflict of Interest

The authors declared no conflict of interest.

Data Availability

Data will be available on request.

Ethical Approval

The study protocol was approved by the Foundation Research Forum Institutional Review Board (IRB No. 2025/056).

REFERENCES

1. Dawson AF, Brown WW, Anderson J, Datta B, Donald JN, Hong K, et al. Mindfulness-based interventions for university students: a systematic review and meta-analysis of randomised controlled trials. Appl Psychol Health Well Being. 2020 Jul;12(2):384-410. https://doi.org/10.1111/aphw.12188
2. Bajaj B, Khoury B, Sengupta S. Resilience and stress as mediators in the relationship of mindfulness and happiness. Front Psychol. 2022 Feb 3;13:771263. https://doi.org/10.3389/fpsyg.2022.771263
3. Sumantry D, Stewart KE. Meditation, mindfulness, and attention: a meta-analysis. Mindfulness. 2021 Jun;12(6):1332-49. https://doi.org/10.1007/s12671-021-01593-w
4. Uddin LQ. Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. Nat Rev Neurosci. 2021 Mar;22(3):167-79. https://doi.org/10.1038/s41583-021-00428-w
5. Sleimen-Malkoun R, Devillers-Réolon L, Temprado JJ. A single session of mindfulness meditation may acutely enhance cognitive performance regardless of meditation experience. PLoS One. 2023 Mar 15;18(3):e0282188. https://doi.org/10.1371/journal.pone.0282188
6. Elcin D, Velasquez M, Colombo PJ. Effects of acute and long-term mindfulness on neural activity and the conflict resolution component of attention. Front Hum Neurosci. 2024 Feb 21;18:1359198. https://doi.org/10.3389/fnhum.2024.1359198
7. Delaquis CP, Joyce KM, Zalewski M, Katz LY, Sulymka J, Agostinho T, et al. Dialectical behaviour therapy skills training groups for common mental health disorders: a systematic review and meta-analysis. J Affect Disord. 2022 Mar 1;300:305-13. https://doi.org/10.1016/j.jad.2021.12.062
8. Villalongo Andino M, Garcia KM, Richey JA. Can dialectical behavior therapy skills group treat social anxiety disorder? A brief integrative review. Front Psychol. 2024 Jan 8;14:1331200. https://doi.org/10.3389/fpsyg.2023.1331200
9. Cavicchioli M, Movalli M, Bruni A, Terragni R, Elena GM, Borgia E, et al. The initial efficacy of stand-alone DBT skills training for treating impulsivity among individuals with alcohol and other substance use disorders. Behav Ther. 2023 Sep 1;54(5):809-22. https://doi.org/10.1016/j.beth.2023.02.006
10. Warner N, Murphy M. Dialectical behaviour therapy skills training for individuals with substance use disorder: a systematic review. Drug Alcohol Rev. 2022 Feb;41(2):501-16. https://doi.org/10.1111/dar.13362
11. Devillers-Réolon L, Mascret N, Sleimen-Malkoun R. Online mindfulness intervention, mental health and attentional abilities: a randomized controlled trial in university students during COVID-19 lockdown. Front Psychol. 2022 Jul 7;13:889807. https://doi.org/10.3389/fpsyg.2022.889807
12. Liu YL, Lee CH, Wu LM. A mindfulness-based intervention improves perceived stress and mindfulness in university nursing students: a quasi-experimental study. Sci Rep. 2024 Jun 8;14(1):13220. https://doi.org/10.1038/s41598-024-64183-5
13. Zhong SY, Guo JH, Zhou XN, Liu JL, Jiang CL. Effects of brief mindfulness meditation training on attention and dispositional mindfulness in young adult males. Acta Psychol (Amst). 2024 Jun 1;246:104277. https://doi.org/10.1016/j.actpsy.2024.104277
14. Thomson JS, Van Hedger SC. Influence of a brief online mindfulness intervention on metacognition, cognition, and emotional outcomes among university students: a randomized longitudinal trial. Psychol Int. 2025 Jul 9;7(3):60. https://doi.org/10.3390/psycholint7030060
15. Jong FJ, Ng AL, Phang CK, Omran S, Teoh SL. The effects of mindfulness-based intervention on cognitive functions in tertiary students: a systematic review and meta-analysis. Appl Cogn Psychol. 2025 May;39(3):e70056. https://doi.org/10.1002/acp.70056
16. Sarfraz A, Siddiqui S, Galante J, Sikander S. Feasibility and acceptability of an online mindfulness-based intervention for stress reduction and psychological wellbeing of university students in Pakistan: a pilot randomized controlled trial. Int J Environ Res Public Health. 2023 Apr 14;20(8):5512. https://doi.org/10.3390/ijerph20085512
17. Reangsing C, Abdullahi SG, Schneider JK. Effects of online mindfulness-based interventions on depressive symptoms in college and university students: a systematic review and meta-analysis. J Integr Complement Med. 2023 May;29(5):292-302. https://doi.org/10.1089/jicm.2022.0606
18. Chen B, Yang T, Xiao L, Xu C, Zhu C. Effects of mobile mindfulness meditation on the mental health of university students: systematic review and meta-analysis. J Med Internet Res. 2023 Jan 3;25:e39128. https://doi.org/10.2196/39128
19. Küchler AM, Schultchen D, Dretzler T, Moshagen M, Ebert DD, Baumeister H. A three-armed randomized controlled trial to evaluate the effectiveness, acceptance, and negative effects of StudiCare mindfulness, an internet- and mobile-based intervention for college students with no and "on demand" guidance. Int J Environ Res Public Health. 2023 Feb 11;20(4):3208. https://doi.org/10.3390/ijerph20043208
20. El Morr C, Ritvo P, Ahmad F, Moineddin R; MVC Team. Effectiveness of an 8-week web-based mindfulness virtual community intervention for university students on symptoms of stress, anxiety, and depression: randomized controlled trial. JMIR Ment Health. 2020 Jul 17;7(7):e18595. https://doi.org/10.2196/18595
21. Ritvo P, Ahmad F, El Morr C, Pirbaglou M, Moineddin R; MVC Team. A mindfulness-based intervention for student depression, anxiety, and stress: randomized

controlled trial. JMIR Ment Health. 2021 Jan 11;8(1):e23491. https://doi.org/10.2196/23491
22. Yamaya N, Ikeda S, Hoshino Y, Takeuchi H, Kawashima R. Temporal changes in the state effect of meditation on response inhibition processes. Mindfulness. 2023 Mar;14(3):554-64. https://doi.org/10.1007/s12671-022-02064-6
23. Devillers-Réolon L, Temprado JJ, Sleimen-Malkoun R. Mindfulness meditation and bimanual coordination control: study of acute effects and the mediating role of cognition. Front Psychol. 2023 May 15;14:1162390. https://doi.org/10.3389/fpsyg.2023.1162390
24. Qomi NJ, Sharifi T, Zadeh SM, Ahmadi R. Examining th efficacy of mindfulness-based stress reduction training on experiential avoidance and impulsivity in female students with difficulty in emotional regulation. J Appl Psychol Res. 2024;15(1):149-67. https://doi.org/10.22059/JAPR.2024.330421.643988
25. Baumgartner JN, Schneider TR. A randomized controlled trial of mindfulness-based stress reduction on academic resilience and performance in college students. J Am Coll Health. 2023 Jul 24;71(6):1916-25. https://doi.org/10.1080/07448481.2021.1950728

Original Article

Allied Medical Research Journal
Vol:4 No.3: 45-55
© The Author (s) 2026

Comparative Effects of Dry Needling and Kinesio Taping as Adjuncts to Strength Training on Pain, Functional Mobility, and Quality of Life in Patients with Knee Osteoarthritis: A Three-Arm Randomized Controlled Trial

Hifza Imtiaz Ahmed[1], Zainab Azhar[1], Mahnoor Mahmood Akhtar[2], Aleeza Akhtar[3], Wafa zainab[3], Zainab Gulshad[3]

[1]*Mukabbir University of Science and Technology, Gujrat, Pakistan*
[2]*University of Sargodha, Sargodha, Pakistan*
[3]*Physio Care Center, Gujrat, Pakistan*

ABSTRACT

Background: Knee osteoarthritis (OA) is the most prevalent form of OA contributing to around 60% of all the cases globally. Strength training is the recommended first-line non-drug treatment, but evidence comparing it with adjunctive therapies such as dry needling and Kinesio taping is yet to be explored.

Methods: This study was a single-masked, three-arm RCT conducted at the Physio Care Center, Gujrat, from November 2024 to December 2025. Fifty participants aged 45 to 65 with knee OA and a Visual Analog Scale pain score ≥7 were randomly assigned to one of three groups: dry needling plus strength training (n=17), Kinesio taping plus strength training (n=17), or strength training alone (n=16), all conducted three times weekly for six weeks. Pain, mobility, and quality of life were assessed at baseline and week 6 by a blinded assessor.

Results: All groups showed significant improvements ($p<0.001$). One-way ANOVA revealed significant differences among groups for pain ($p=0.010$), mobility ($p=0.003$), and quality of life ($p<0.001$). Dry needling plus strength training resulted in significantly greater improvements than strength training alone ($p\leq0.006$), whereas Kinesio taping showed no significant advantage over strength training alone.

Conclusion: Adding dry needling to strength training yields significantly greater short-term gains in pain, mobility, and quality of life than strength training alone in knee osteoarthritis, whereas the added benefit of Kinesio taping is smaller and less certain. Larger, sham-controlled, multicenter trials with longer follow-up are warranted.

Keywords: *Athletic Tape; Dry Needling; Knee; Osteoarthritis; Quality of Life; Resistance Training.*

Received: April 17th, 2026**; Revised:** June 20th, 2026**; Accepted:** July 12th, 2026

Corresponding Email: hefzawarraich286@gmail.com

DOI: https://doi.org/10.59564/amrj/04.03/006

INTRODUCTION

Knee Osteoarthritis (OA) is a chronic progressive degenerative joint disease and is the most prevalent form of OA in the world, with about 60% of all reported cases of OA occurring in the knee[1-2].It is defined by cartilage degradation, sub-chondral bone alterations, osteophyte growth, synovial inflammation, and progressive muscle weakness, resulting in chronic pain, stiffness, limited motion, and functional disability[3,4]. The pathological changes significantly diminish patients' quality of life, interfere with participation in daily activities and occupational roles, and create a significant economic burden due to utilization of healthcare resources and loss of

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

productivity[5]. Current approaches to treatment, such as simple care, pharmacologic treatment, and surgery, may be only temporarily effective and may carry risks of adverse effects, high costs, or surgical complications. With no cure, current recommendations focus on non-pharmacological, low-risk therapies to manage pain, maintain joint function, and slow disease progression[6,7,8].

Exercise therapy, especially strength training, is a widely recommended first-line non-pharmacological treatment for knee OA, given the strong association between quadriceps weakness and disease progression[9]. Strengthening the muscles around the knee will increase the joint's stability, shock absorption, and load distribution, and several randomized trials have shown short-term increases in muscle strength and cardiorespiratory fitness after structured strength or aerobic exercise programs[3]. The long-term effectiveness of exercise alone on quality of life, however, is not consistent, with some trials reporting only small or statistically non-significant improvements at longer-term follow-up, possibly due to small sample sizes, heterogeneous patient populations, and reduced adherence over time[3]. This has highlighted the need for additional therapies to augment, expedite, or sustain the positive effects of exercise-based rehab.

Acupuncture-related methods, such as dry needling of myofascial trigger points and ashi points on the quadriceps, have demonstrated efficacy in reducing pain and improving function, and are thought to work via modulation of nociceptive pathways at the spinal and central levels, endogenous release of analgesic substances, and improved local muscle relaxation and blood flow. Synergistic effects have been observed in trials in which acupuncture was used in conjunction with exercise rehabilitation, resulting in better, faster symptom relief than either treatment alone, and the combined groups had higher response rates than exercise alone. Likewise, manual therapy with neuromuscular training is more effective than traditional electrotherapy-based interventions in reducing pain, increasing knee flexion range of motion, balance, and quality of life over 6 weeks of intervention[10]. Another emerging technique that has gained popularity is Kinesio taping, which is believed to increase the torque of the quadriceps, decrease pain, and stimulate the cutaneous and mechanoreceptors around the knee joint, and has been utilized by clinicians for its non-invasive, easy-to-apply nature[10,11]. The related non-invasive technique of acupressure has also shown considerable benefits for pain, stiffness, physical function, and health-related quality of life compared to routine care[12].

Although individual studies have shown positive effects of both dry needling and Kinesio taping as adjuncts to exercise, direct comparisons between the two modalities, in conjunction with a standardized strength training regimen, remain limited. Existing trials have tested the effect of acupuncture, manual therapy, or taping alone against usual care, sham treatments, or electrotherapy, but not against each other in the same well-designed trial. This results in clinicians lacking definitive comparative evidence to inform their choice of the most effective adjunct therapy to accompany strength training for managing knee OA. A randomized controlled trial directly comparing these interventions is therefore warranted to guide clinical decision-making regarding evidence-based rehabilitation protocols for this very common and debilitating condition. This 3-arm randomized controlled trial aims to assess the efficacy of dry needling, Kinesio taping, and strength training compared to strength training alone for the treatment of knee osteoarthritis on pain, functional mobility, and quality of life to determine if either of these adjunctive therapies is superior to strength training alone, and if either is superior.

METHODOLOGY

Study Design and Setting

Study Design was three-arm, parallel, randomized controlled trial (RCT) with single blinded assessment. It was conducted at the Outpatient Physiotherapy Department of the Physio Care Center, Gujarat from November 2024 to December 2025.

Selection Criteria

Participants were recruited from the outpatient physiotherapy and orthopedic department. Eligible participants meeting the inclusion criteria were invited to participate after a detailed explanation of the study protocol, following which written informed consent was obtained.

***Inclusion Criteria*:** was based on following criteria (i) clinically confirmed diagnosis of knee osteoarthritis; (ii) were aged between 45 and 65

years and (iii) reported a Visual Analogue Scale (VAS) pain score of 7 or above.

***Exclusion Criteria*:** Participants were excluded if they (i) had any known allergic conditions, including sensitivity to adhesive tape or needling materials (ii) recent knee surgery and (iii) unfit to perform strength training exercises of lower limb

Sample Size

G*Power software was used to determine a sample size, based on the mean difference of 1.7 in the Visual Analogue Scale (VAS) pain score reported between intervention groups following comparable physiotherapy interventions in patients with knee osteoarthritis, calculated at a 95% confidence interval[13]. A sample size of n=50 was estimated using the following formula for comparison of means between independent groups:

$$n = 2\sigma^2(Z_{\alpha/2} + Z_{\beta})^2 / d^2$$

Randomization and Blinding

Eligible participants were randomly allocated to one of the three intervention groups using the lottery (chit) method of randomization. Fifty numbered chits, corresponding to Group A (Dry Needling with Strength Training), Group B (Kinesio Taping with Strength Training), and Group C (Strength Training alone), were prepared in equal proportion, folded, and placed in a sealed, opaque, non-transparent container by an independent investigator who was not involved in participant recruitment, treatment delivery, or outcome assessment. At the time of enrolment, each eligible participant was asked to draw one chit from the container, and the group written on the drawn chit determined the participant's allocation.

Intervention Protocol

Each group was assigned an intervention delivered three times a week for six weeks.

- **Group A: Dry Needling Combined with Strength Training**
  Participants were treated with dry needling of myofascial trigger points in the quadriceps, specifically the Vastus Medialis Obliquus, and periarticular musculature of the affected knee with sterile filiform needles inserted at a depth to produce a local twitch response and then a standardized strength training program.

- **Group B: Kinesio Taping Combined with Strength Training**
  Participants were given Kinesio taping over the quadriceps muscle using the muscle facilitation technique and around the patella using the patellar correction technique for pain reduction, increased proprioceptive feedback, and increased quadriceps activation, in addition to the same strength training program.

- **Group C: Strength Training**
  Group C was given the strength training program only, which involved quadriceps-setting exercises, straight-leg raises, mini squats, and resisted knee extension exercises, all performed in 2-3 sets of 10-12 repetitions, with sets and repetitions increased as tolerated by the subject.

Outcome Measures

A trained, blinded assessor collected all outcome measures at baseline (Week 0) and post-intervention (Week 6).

i. **Numeric Pain Rating Scale (NPRS)** (0 = no pain, 10 = worst imaginable pain) was used to assess pain intensity. The NPRS is a valid, reliable, and responsive instrument commonly used to measure pain intensity in persons with knee OA and is responsive to changes in physiotherapy treatment[14].

ii. **Timed Up and Go (TUG) test** was used to measure Functional mobiity, which requires the participant to get up from a standard armchair, walk 3 meters, turn, walk back, and sit down again, and the time taken (in seconds) was recorded. The TUG is highly reliable and valid in assessing functional mobility in patients with knee OA[15].

iii. **The World Health Organization Quality of Life-BREF (WHOQOL-BREF)** is a 26-item validated questionnaire that measures quality of life in four domains: physical health, psychological health, social relationships, and environment, with higher scores indicating better quality of life. WHOQOL-BREF has been tested in various populations of people with chronic musculoskeletal conditions, including knee osteoarthritis[16].

Statistical Analysis

Data were entered and analyzed in SPSS version 26 (IBM Corp., USA). Skewness and Kurtosis were used to determine the normality of the data distribution. Baseline demographic and clinical data were summarized using descriptive statistics, including means, standard deviations, frequencies, and percentages. Repeated-measures analysis was used to compare NPRS, TUG, and WHOQOL-BREF scores within groups from baseline to post-intervention and between groups across the three arms using one-way analysis of variance (ANOVA), followed by pairwise comparisons with Bonferroni correction. The $p<0.05$ was considered statistically significant.

Ethical Consideration

The Institutional Review Board approved the study protocol (FMRL-IRB/2024/026) before the study began, and the trial was registered in advance on ClinicalTrials.gov (Identifier: NCT07710235). All participants received comprehensive information about the study procedures, risks, and benefits, and signed informed consent before enrollment. The study was carried out in compliance with the ethical principles of the Declaration of Helsinki.

RESULTS

Demographic Characteristics

Fifty participants were enrolled and randomly divided into 3 groups: Group A (Dry Needling + Strength Training, n = 17); Group B (Kinesio Taping + Strength Training, n = 17); Group C (Strength Training alone, n = 16). There were no dropouts in the 6-week intervention period, and all 50 participants completed baseline and post-intervention assessments (Figure 1).

The overall mean age of participants was 54.94 ± 7.19 years, and the mean BMI was 27.28 ± 3.38 kg/m². There were no statistically significant differences between the three groups at baseline, with respect to age, BMI, symptom duration, gender distribution, Kellgren–Lawrence (KL) radiographic grade, or affected side (all $p > 0.05$), indicating successful randomization (Table 1).

Table 1. Baseline Demographic and Clinical Characteristics of Participants

Variable	Group A (n=17)	Group B (n=17)	Group C (n=16)	p-value
Age (years)	55.41 ± 6.41	56.35 ± 7.36	52.94 ± 7.78	0.381
BMI (kg/m²)	27.46 ± 4.01	26.17 ± 2.73	28.25 ± 3.12	0.204
Symptom duration (months)	17.47 ± 7.00	17.76 ± 7.85	19.50 ± 7.14	0.697
Gender, Female n (%)	14 (82.4)	12 (70.6)	11 (68.8)	0.622
Gender, Male n (%)	3 (17.6)	5 (29.4)	5 (31.3)	—
KL Grade II, n (%)	12 (70.6)	11 (64.7)	9 (56.3)	0.690
KL Grade III, n (%)	5 (29.4)	6 (35.3)	7 (43.8)	—
Affected side: Right, n (%)	7 (41.2)	6 (35.3)	5 (31.3)	0.964
Affected side: Left, n (%)	7 (41.2)	8 (47.1)	7 (43.8)	—
Affected side: Bilateral, n (%)	3 (17.6)	3 (17.6)	4 (25.0)	—

Group A = Dry Needling + Strength Training; Group B = Kinesio Taping + Strength Training; Group C = Strength Training alone; BMI = Body Mass Index; KL = Kellgren–Lawrence. Continuous variables analyzed by one-way ANOVA; categorical variables analyzed by Chi-square test. Mean ± SD; n (%).

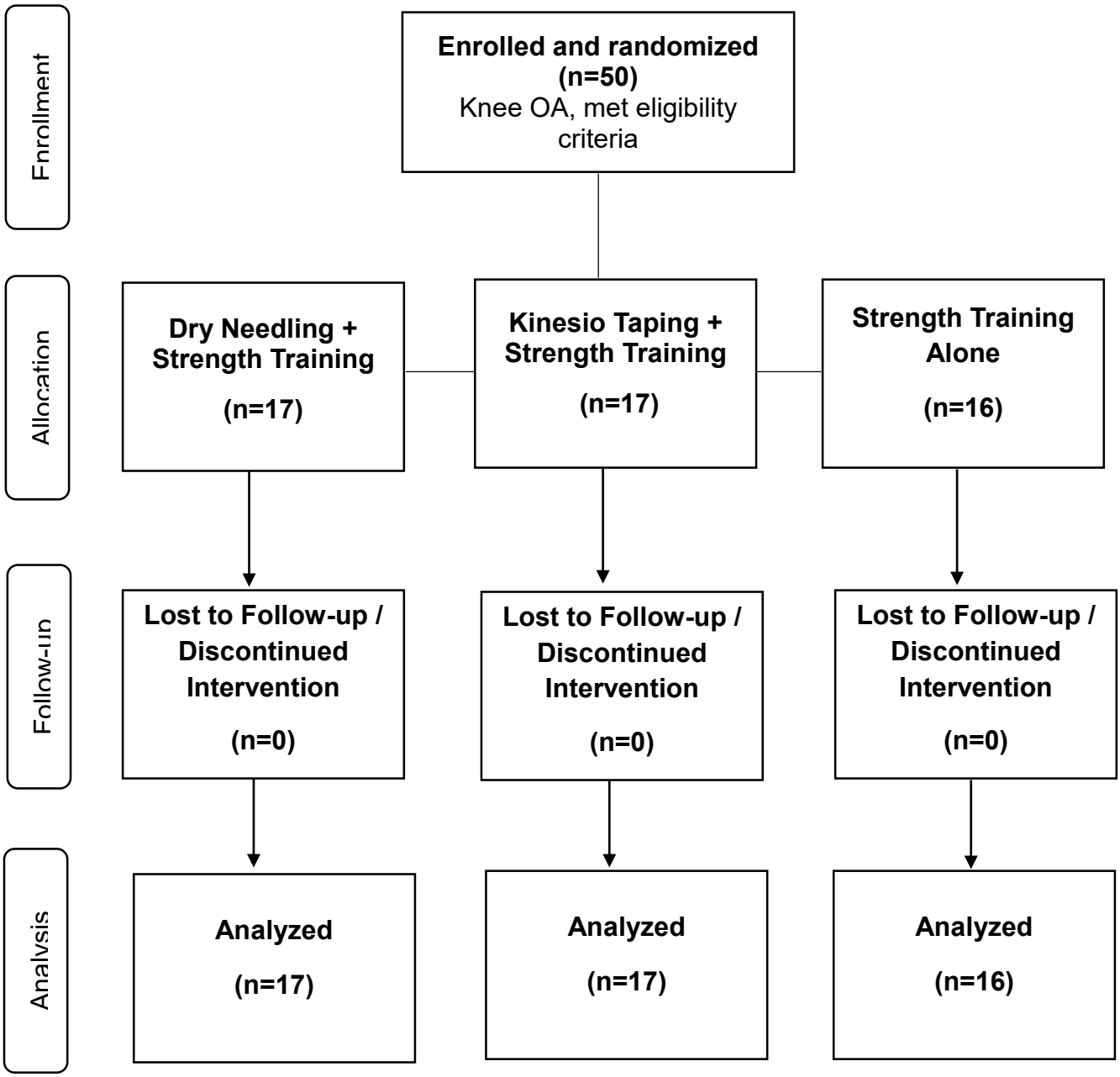

Figure 1. CONSORT flow diagram of participant enrollment, allocation, follow-up, and analysis

Adverse Events

There were no serious adverse events in any of the three groups in the six weeks of the intervention. A small number of participants in Group A reported soreness after the needle insertions and minor bruising at the sites where the needles were inserted. In contrast, a small number of participants in Group B were slightly irritated by the Kinesio tape; this irritation was minor, self-limited, and managed with conservative treatment by the treating physiotherapist; no participant required treatment

modification, medical referral, or withdrawal from the study as a result.

Normality Testing

All baseline and post-intervention outcome variables were subjected to the Shapiro-Wilk test. Overall scores on the WHOQOL-BREF were normally distributed at baseline and post-intervention ($p > 0.05$). In contrast, NPRS and TUG scores showed minor deviations from normality at specific time points, likely due to the discrete, narrow range of the NPRS scale. All outcome measures were analyzed using parametric tests as outlined a priori, namely, a paired-samples t-test for within-group comparisons and a one-way ANOVA with Bonferroni-adjusted post hoc pairwise comparisons for between-group comparisons, given the approximately symmetric distribution of the data and adequate sample size per group.

Baseline Comparability

One-way ANOVA confirmed that there were no statistically significant differences between the groups at baseline in terms of NPRS ($F = 0.018$, $p = 0.982$), TUG ($F = 0.890$, $p = 0.418$), or WHOQOL-BREF overall scores ($F = 0.385$, $p = 0.682$), indicating that the three groups were comparable. Any differences following the intervention were due to the intervention effect.

Within-Group Effects after 6 Weeks of Treatment

Within-group difference scores were statistically significant for all three groups for all outcome measures after the 6-week intervention period (all $p < 0.001$) (Table 2). Group A showed the largest reduction in pain, with NPRS decreasing from 7.12 ± 1.05 to 3.94 ± 1.68 ($t = 11.579$, $p < 0.001$), while Group B improved from 7.12 ± 0.99 to 4.53 ± 1.33 ($t = 10.029$, $p < 0.001$), and Group C improved from 7.06 ± 0.77 to 5.62 ± 1.59 ($t = 5.578$, $p < 0.001$).

Functional mobility, measured by the TUG test, improved significantly in all groups, with Group A showing the greatest reduction in completion time (14.02 ± 2.23 to 10.91 ± 2.22 seconds; $t = 17.097$, $p < 0.001$), followed by Group B (14.87 ± 1.93 to 12.72 ± 2.38 seconds; $t = 9.279$, $p < 0.001$) and Group C (14.80 ± 1.97 to 13.61 ± 2.01 seconds; $t = 5.694$, $p < 0.001$).

Quality of life, assessed using the WHOQOL-BREF overall score, also improved significantly in all groups, with Group A demonstrating the greatest gain (51.46 ± 4.95 to 62.99 ± 5.77; $t = -30.349$, $p < 0.001$), compared to Group B (49.88 ± 5.36 to 58.09 ± 5.95; $t = -28.913$, $p < 0.001$) and Group C (50.53 ± 5.54 to 55.04 ± 5.20; $t = -10.571$, $p < 0.001$).

Table 2. Within-Group Comparison of Outcome Measures (Baseline vs. Post-6-Week Intervention)

Outcome Measure	Group	Pre Mean ± SD	Post Mean ± SD	t-value	p-value
NPRS (0–10)	A	7.12 ± 1.05	3.94 ± 1.68	11.579	<0.001
	B	7.12 ± 0.99	4.53 ± 1.33	10.029	<0.001
	C	7.06 ± 0.77	5.62 ± 1.59	5.578	<0.001
TUG (sec)	A	14.02 ± 2.23	10.91 ± 2.22	17.097	<0.001
	B	14.87 ± 1.93	12.72 ± 2.38	9.279	<0.001
	C	14.80 ± 1.97	13.61 ± 2.01	5.694	<0.001
WHOQOL-BREF Overall (0–100)	A	51.46 ± 4.95	62.99 ± 5.77	−30.349	<0.001
	B	49.88 ± 5.36	58.09 ± 5.95	−28.913	<0.001
	C	50.53 ± 5.54	55.04 ± 5.20	−10.571	<0.001

NPRS = Numeric Pain Rating Scale; TUG = Timed Up and Go test; WHOQOL-BREF = World Health Organization Quality of Life-BREF. Within-group comparison by paired samples t-test. Significance set at $p < 0.05$.

Between-Group Comparison after 6 Weeks of Treatment

Statistically significant between-group differences were found at post-intervention for all three primary outcome measures: NPRS ($F = 5.075$, $p = 0.010$), TUG ($F = 6.450$, $p = 0.003$), and WHOQOL-BREF overall score ($F = 8.338$, $p < 0.001$) (Table 3).

Post-hoc pairwise comparisons with Bonferroni correction (adjusted $\alpha = 0.017$) showed that Group A (Dry Needling + Strength Training) performed significantly better than Group C (Strength Training alone), in all three measures (NPRS, $p = 0.006$; TUG, $p < 0.001$; WHOQOL-BREF overall, $p < 0.001$). Group A also showed greater improvement than Group B across all outcomes.

However, this was not significant after Bonferroni correction for NPRS ($p = 0.265$), TUG ($p = 0.028$), or WHOQOL-BREF overall ($p = 0.020$). After correction, there was no statistically significant difference between the mean scores of Group B and Group C for NPRS ($p = 0.039$) or TUG ($p = 0.254$); however, Group B consistently had higher mean scores than Group C for all outcomes.

Table 3. Between-Group Comparison of Post-Intervention Outcome Measures

Outcome Measure	Group A (n=17)	Group B (n=17)	Group C (n=16)	F-value	p-value
NPRS (0–10)	3.94 ± 1.68	4.53 ± 1.33	5.62 ± 1.59	5.075	0.010*
TUG (sec)	10.91 ± 2.22	12.72 ± 2.38	13.61 ± 2.01	6.450	0.003*
WHOQOL-BREF Overall (0–100)	62.99 ± 5.77	58.09 ± 5.95	55.04 ± 5.20	8.338	<0.001*

**p < 0.05 (one-way ANOVA). Between-group comparison by one-way ANOVA.*

Table 4. Post-Hoc Pairwise Comparisons (Bonferroni-Adjusted, α = 0.017)

Outcome Measure	Comparison	Mean Difference	t-value	p-value	Significance
NPRS	A vs. B	−0.59	−1.134	0.265	NS
	A vs. C	−1.68	−2.960	0.006	Significant
	B vs. C	−1.09	−2.156	0.039	NS
TUG	A vs. B	−1.81	−2.295	0.028	NS
	A vs. C	−2.70	−3.662	<0.001	Significant
	B vs. C	−0.89	−1.163	0.254	NS
WHOQOL-BREF Overall	A vs. B	4.90	2.442	0.020	NS
	A vs. C	7.95	4.150	<0.001	Significant
	B vs. C	3.05	1.562	0.129	NS

NS = not significant after Bonferroni correction (adjusted α = 0.017).

WHOQOL-BREF Domain-Specific Analysis

Analysis of the individual WHOQOL-BREF domains revealed significant within-group improvements across all four domains (Physical, Psychological, Social Relationships, and Environmental) in all three groups (all $p < 0.001$) (Table 5).

Between-group analysis at post-intervention showed a statistically significant difference only in the Environmental domain ($F = 3.548$, $p = 0.037$), with Group A demonstrating the greatest improvement (52.89 ± 8.23 to 66.09 ± 9.60).

Differences between groups did not reach statistical significance for the Physical ($F = 1.728$, $p = 0.189$), Psychological ($F = 2.378$, $p = 0.104$), or Social Relationships ($F = 2.893$, $p = 0.065$) domains, although Group A consistently showed numerically greater gains across all four domains compared to Groups B and C.

Table 5. Within-Group Comparison of WHOQOL-BREF Domain Scores

Domain	Group	Pre Mean ± SD	Post Mean ± SD	p-value	Between-Group p
Physical Health	A	49.31 ± 10.76	60.76 ± 11.96	<0.001	0.189
	B	48.76 ± 9.83	57.68 ± 10.89	<0.001	
	C	49.19 ± 7.82	54.12 ± 7.05	<0.001	
Psychological	A	51.26 ± 7.54	61.40 ± 10.39	<0.001	0.104
	B	48.17 ± 9.66	56.84 ± 10.66	<0.001	
	C	48.44 ± 8.83	53.54 ± 10.18	<0.001	
Social Relationships	A	52.41 ± 9.49	63.75 ± 9.42	<0.001	0.065
	B	49.65 ± 11.16	56.19 ± 11.33	<0.001	
	C	52.08 ± 10.39	55.80 ± 11.57	<0.001	
Environmental	A	52.89 ± 8.23	66.09 ± 9.60	<0.001	0.037*
	B	52.94 ± 9.93	61.64 ± 11.30	<0.001	
	C	52.39 ± 9.04	56.76 ± 9.07	<0.001	

**Statistically significant at p < 0.05. Within-group comparison by paired samples t-test; between-group comparison by one-way ANOVA*

DISCUSSION

This three-arm randomized controlled trial compared dry needling combined with strength training, Kinesio taping combined with strength training, and strength training alone in patients with knee osteoarthritis. All three groups achieved statistically significant within-group improvements in pain, functional mobility, and quality of life after six weeks, confirming that structured strength training itself is an effective, first-line intervention for this population[17]. However, the addition of dry needling conferred a significantly greater benefit than strength training alone across all three primary outcomes, while the addition of Kinesio taping produced a consistent, but statistically non-significant, numerical advantage over strength training alone. Dry needling combined with strength training also showed a numerically greater improvement than Kinesio taping combined with strength training on every outcome, although this difference did not survive Bonferroni correction.

The superior performance of dry needling combined with strength training is consistent with a growing body of trial evidence showing that needling directed at myofascial trigger points and periarticular musculature produces meaningful reductions in pain and disability when layered onto exercise-based rehabilitation for knee osteoarthritis[18,19,20,21,22]. Ma et al. reported that dry needling of latent and active trigger points produced pain and functional improvements comparable to, and in some respects exceeding, oral diclofenac when combined with stretching[19], while Dunning et al. demonstrated in a large multicenter trial that periosteal electrical dry needling added to manual therapy and exercise produced significantly greater reductions in disability and pain than manual therapy and exercise alone, with benefits sustained to three months[20]. Similarly, Pang et al. found that both ultrasound-guided and traditional dry needling combined with exercise therapy outperformed exercise therapy alone for pain relief[21]. The proposed mechanisms underlying these effects include mechanical disruption of dysfunctional motor endplates and taut bands within trigger points, modulation of nociceptive signalling at the spinal and supraspinal level, release of endogenous opioids and other analgesic substances, and improved local muscle blood flow and relaxation[18,23,24], all of which may act synergistically with the biomechanical benefits of strength training to accelerate and augment recovery.

The evidence supporting the use of Kinesio taping as a complementary therapy to strengthening exercises for knee OA, on the other hand, is far less consistent. In Group B, this effect was statistically significant in only 1 of 12 outcomes, although it was numerically consistent across all outcomes. Several trials and a meta-analysis have reported short-term reductions in pain and improvements in function with the addition of Kinesio taping or when compared with sham taping, which they have attributed to increased quadriceps activation, stimulation of the skin and mechanoreceptors, and better proprioceptive feedback around the patellofemoral joint[10,11,25,26]. Other trials, however, have shown no meaningful differences between Kinesio taping and sham taping or usual care on pain, strength, or function[27], and a systematic review, limited to Kinesio taping with concurrent physical therapy, found that the effect of Kinesio taping on pain was small and short-lived[27]. The present findings, which did not reach significance when compared with strength training alone but trended toward benefit, fall comfortably within this disparate literature and suggest that the benefits of Kinesio taping may be a smaller and less consistent addition to exercise than those of dry needling.

This interpretation is supported by the pattern of results obtained on the three functional mobility and quality-of-life measures. The greater reduction in Timed Up corroborates the increase in the strength of the three arms and Go completion time in the dry needling group, consistent with more effective desensitization of periarticular trigger points, allowing for fuller recruitment of the quadriceps during the sit-to-stand and gait components of the task. Similarly, the only quality-of-life domain to show a significant difference between groups in this study was the Environmental domain. In contrast, the other two domains (Physical and Psychological/Social Relationships) showed improvements across all treatment arms, indicating that the most immediate cross-domain benefit of superior pain and mobility control in the 6 weeks of this study is the ability to negotiate one's immediate environment and daily surroundings. In contrast, broader psychological and social gains may require a longer intervention or follow-up period to consolidate and differentiate between treatment arms. This is consistent with findings from other combined physiotherapy trials in knee OA, where combining active exercise with adjunct therapy showed better and more rapid symptomatic recovery than active exercise or comparators (including electrotherapy) alone[18], and confirms that exercise therapy is the mainstay of non-surgical knee OA treatment regardless of the adjunct used, which enhances and accelerates the benefits of exercise therapy alone[2,28].

Strengths and Limitations

There are several strengths of this trial. The three-arm design allowed for a direct comparison between two emerging therapies and a common, standardized strength training intervention, which is not often seen in the current research literature, and which fills an important research gap. There was excellent baseline comparability between groups, with no significant differences in age, BMI, symptom duration, sex distribution, radiographic severity, or affected side. A blinded assessor conducted outcome assessment, and the measures used were reliable, valid, and responsive in populations with knee osteoarthritis. There was no attrition bias, as all 50 participants enrolled in the study completed both baseline and post-intervention assessments, which increased confidence in the within-group and between-group findings.

There are, however, some caveats to consider when interpreting these results. The trial was single-blind: outcome assessors were blinded, but participants and treating therapists were not, because the nature of the intervention (needling and taping) made this impossible. This may have led to performance and expectation bias, especially for the self-reported NPRS and WHOQOL-BREF measures. Due to the lack of sham dry needling or sham taping arm, the extra contact time, attention, and therapeutic ritual provided by each adjunct cannot be separated from any modality-specific physiological effect, meaning that the real specific effect of either technique may be smaller than the differences between groups. The non-significant but numerically consistent advantage of dry needling over Kinesio taping after Bonferroni correction may be due to the small number of cases in each arm of this trial (n = 17, 17, and 16), which may have limited statistical power to detect smaller, clinically relevant differences between the two adjunct groups themselves; a larger, multicenter trial would be better powered to resolve this comparison. The durability of the observed benefits and whether the difference between groups increases, decreases, or disappears over longer follow-up were not assessed. Lastly, while

this trial did not report any serious adverse events, the minor procedure-related adverse events that did occur were treated conservatively by the treating physiotherapist. Still, the trial did not employ a standardized instrument to grade adverse events or treatment tolerability systematically, nor was cost assessed, both of which are relevant to clinicians and patients choosing between adjunct therapies in routine practice.

CONCLUSION

After six weeks, when compared to strength training alone, the combination of dry needling and a structured strength training program resulted in significantly greater improvements in pain, functional mobility, and quality of life in patients with knee osteoarthritis, while the combination of Kinesio taping and strength training resulted in a consistent but statistically non-significant numerical advantage compared to strength training alone. Dry needling also demonstrated a greater numeric benefit, though not statistically significant, compared to Kinesio taping across all outcomes. The results indicate that dry needling is a clinically useful, low-risk adjunct that physiotherapists can use in strength-training programs to improve short-term outcomes in knee osteoarthritis. In contrast, Kinesio taping's incremental effect is less clear and smaller. These results support the need for larger, multicenter, sham-controlled trials with longer-term follow-up to confirm the findings, better understand the durability of treatment effects, and determine the comparative cost-effectiveness of these adjunct therapies in routine clinical practice.

Author Contributions

HIA: Conception and Design, Supervision, and Manuscript Writing.
ZA: Study Design, Data Collection, and Data Analysis.
MMA: Literature Review, Data Interpretation, and Manuscript Preparation.
AA & WZ: Data Collection, Data Entry, and Manuscript Editing.
ZG: Literature Review, Critical Review, and Final Approval.

Funding Source

No sources.

Conflict of Interest

The authors declared no conflict of interest.

Data Availability

Data will be available on request.

Ethical Approval

The Institutional Review Board approved the study protocol (FMRL-IRB/2024/026) before the study began, and the trial was registered in advance on ClinicalTrials.gov (Identifier: NCT07710235).

REFERENCES

1. Long H, Liu Q, Yin H, Wang K, Diao N, Zhang Y, Lin J, Guo A. Prevalence trends of site-specific osteoarthritis from 1990 to 2019: findings from the global burden of disease study 2019. Arthritis & Rheumatology. 2022 Jul;74(7):1172-83. doi: https://doi.org/10.1002/art.42089
2. Sharma L. Osteoarthritis of the knee. New England Journal of Medicine. 2021 Jan 7;384(1):51-9. doi: https://doi.org/10.1056/NEJMcp1903768
3. Øiestad BE, Årøen A, Røtterud JH, Østerås N, Jarstad E, Grotle M, Risberg MA. The efficacy of strength or aerobic exercise on quality of life and knee function in patients with knee osteoarthritis. A multi-arm randomized controlled trial with 1-year follow-up. BMC Musculoskeletal Disorders. 2023 Sep 8;24(1):714. doi: https://doi.org/10.1186/s12891-023-06831-x
4. Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: a review. JAMA. 2021 Feb 9;325(6):568-79. doi: https://doi.org/10.1001/jama.2020.22171
5. Aydemir T, Çirpan R, Kartin PT. Impact of acupressure on pain, functional status, and quality of life in individuals with knee osteoarthritis. Journal of Manipulative and Physiological Therapeutics. 2025 Jan;48(1-5):204-13. doi: https://doi.org/10.1016/j.jmpt.2025.09.001
6. Li H, Qian S, Liu Y, Guo Y, Chen X, Hu Y, Fu Q, Li Q, Xiang Y, Liu Y, Wu X, Luo T. Opposing needling in the treatment of knee osteoarthritis: an improved study protocol for a randomized controlled trial. Frontiers in Medicine. 2025 Nov 25;12:1698983. doi: https://doi.org/10.3389/fmed.2025.1698983
7. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, Callahan L, Copenhaver C, Dodge C, Felson D, Gellar K. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis & Rheumatology. 2020 Feb;72(2):220-33. doi: https://doi.org/10.1002/art.41142
8. Whittaker JL, Truong LK, Dhiman K, Beck C. Osteoarthritis year in review 2020: rehabilitation and outcomes. Osteoarthritis and Cartilage. 2021 Feb;29(2):190-207. doi: https://doi.org/10.1016/j.joca.2020.10.005
9. Skou ST, Roos EM. Physical therapy for patients with knee and hip osteoarthritis: supervised, active treatment is current best practice. Clinical and Experimental Rheumatology. 2020 Sep;38(5):1036-9
10. Mohamed SH, Albalawi HF, Nambi G, Shalabi KM, Albalawi MO. Effectiveness of manual therapy vs conventional physical therapy with neuromuscular training in the management of knee osteoarthritis: a randomized clinical control trial. International Journal of Osteopathic Medicine. 2025 Sep 11:100787. doi: https://doi.org/10.1016/j.ijosm.2025.100787
11. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines.

Osteoarthritis and Cartilage. 2008 Feb;16(2):137-62. doi: https://doi.org/10.1016/j.joca.2007.12.013
12. Holden MA, Hattle M, Runhaar J, Riley RD, Healey EL, Quicke J, van der Windt DA, Dziedzic K, van Middelkoop M, Burke D, Legha A. Moderators of the effect of therapeutic exercise for knee and hip osteoarthritis: a systematic review and individual participant data meta-analysis. The Lancet Rheumatology. 2023 Jul;5(7):e386-400. doi: https://doi.org/10.1016/S2665-9913(23)00122-4
13. Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement properties of visual analogue scale, numeric rating scale, and pain severity subscale of the brief pain inventory in patients with low back pain: a systematic review. The Journal of Pain. 2019 Mar;20(3):245-63. doi: https://doi.org/10.1016/j.jpain.2018.07.009
14. Euasobhon P, Atisook R, Bumrungchatudom K, Zinboonyahgoon N, Saisavoey N, Jensen MP. Reliability and responsivity of pain intensity scales in individuals with chronic pain. Pain. 2022 Dec 1;163(12):e1184-91. doi: https://doi.org/10.1097/j.pain.0000000000002692
15. Svinøy OE, Hilde G, Bergland A, Strand BH. Timed up and go: reference values for community-dwelling older adults with and without arthritis and non-communicable diseases: the Tromsø study. Clinical Interventions in Aging. 2021 Feb 23:335-43. doi: https://doi.org/10.2147/CIA.S294512
16. The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): development and general psychometric properties. Social Science & Medicine. 1998 Jun 15;46(12):1569-85. doi: https://doi.org/10.1016/s0277-9536(98)00009-4
17. Hinman RS, Hall M, Comensoli S, Bennell KL. Exercise and Sports Science Australia (ESSA) updated position statement on exercise and physical activity for people with hip/knee osteoarthritis. Pensar en Movimiento: Revista de Ciencias del Ejercicio y la Salud. 2024 Dec;22(2):228-65. doi: https://doi.org/10.15517/pensarmov.v22i2.62992
18. Liu XY, Ma Y, Huang ZY, Xiao XX, Guan L. The efficacy of acupuncture, exercise rehabilitation, and their combination in the treatment of knee osteoarthritis: a randomized controlled trial. Journal of Pain Research. 2024 Dec 31:2837-49. doi: https://doi.org/10.2147/JPR.S465058
19. a YT, Dong YL, Wang B, Xie WP, Huang QM, Zheng YJ. Dry needling on latent and active myofascial trigger points versus oral diclofenac in patients with knee osteoarthritis: a randomized controlled trial. BMC Musculoskeletal Disorders. 2023 Jan 18;24(1):36. doi: https://doi.org/10.1186/s12891-022-06116-9
20. Young I, Dunning J, Butts R, Bliton P, Zacharko N, Garcia J, Mourad F, Charlebois C, Gorby P, Fernández-de-Las-Peñas C. Spinal manipulation and electrical dry needling as an adjunct to conventional physical therapy in patients with lumbar spinal stenosis: a multi-center randomized clinical trial. The Spine Journal. 2024 Apr;24(4):590-600. doi: https://doi.org/10.1016/j.spinee.2023.12.002
21. Pang JC, Fu AS, Lam SK, Peng B, Fu AC. Ultrasound-guided dry needling versus traditional dry needling for patients with knee osteoarthritis: a double-blind randomized controlled trial. PLoS One. 2022 Sep 30;17(9):e0274990. doi: https://doi.org/10.1371/journal.pone.0274990
22. Hunter CW, Deer TR, Jones MR, Chang Chien GC, D'Souza RS, Davis T, Eldon ER, Esposito MF, Goree JH, Hewan-Lowe L, Maloney JA. Consensus guidelines on interventional therapies for knee pain (STEP guidelines) from the American Society of Pain and Neuroscience. Journal of Pain Research. 2022 Dec 31:2683-745. doi: https://doi.org/10.2147/JPR.S370469
23. Chys M, De Meulemeester K, De Greef I, Murillo C, Kindt W, Kouzouz Y, Lescroart B, Cagnie B. Clinical effectiveness of dry needling in patients with musculoskeletal pain: an umbrella review. Journal of Clinical Medicine. 2023 Feb 2;12(3):1205. doi: https://doi.org/10.3390/jcm12031205
24. De Greef I, Chys M, Gerwin RD, De Meulemeester K, Cagnie B. The neurophysiological effects of dry needling: an update of a narrative review. American Journal of Physical Medicine & Rehabilitation. 2026 Feb 1;105(2):172-80. doi: https://doi.org/10.1097/PHM.0000000000002855
25. Tran L, Makram AM, Makram OM, Elfaituri MK, Morsy S, Ghozy S, Zayan AH, Nam NH, Zaki MM, Allison EL, Hieu TH. Efficacy of kinesio taping compared to other treatment modalities in musculoskeletal disorders: a systematic review and meta-analysis. Research in Sports Medicine. 2023 Jul 4;31(4):416-39. doi: https://doi.org/10.1080/15438627.2021.1989432
26. Mohamed SH, Alatawi SF. Effectiveness of kinesio taping and conventional physical therapy in the management of knee osteoarthritis: a randomized clinical trial. Irish Journal of Medical Science. 2023 Oct;192(5):2223-33. doi: https://doi.org/10.1007/s11845-022-03247-9
27. Mao HY, Hu MT, Yen YY, Lan SJ, Lee SD. Kinesio taping relieves pain and improves isokinetic not isometric muscle strength in patients with knee osteoarthritis: a systematic review and meta-analysis. International Journal of Environmental Research and Public Health. 2021 Oct 4;18(19):10440. doi: https://doi.org/10.3390/ijerph181910440
28. Yao K, Shamim MF, Xia J, Liu TT, Guo Y, Lin X. Therapeutic potential of acupuncture in knee osteoarthritis: clinical efficacy and mechanistic insights. Journal of Inflammation Research. 2025 Dec 31:12169-90. doi: https://doi.org/10.2147/JIR.S526890

Original Article

Allied Medical Research Journal
Vol:4 No.3: 56-64
© The Author (s) 2026

Elemental Profile of Postoperative Infected Wound Exudates in Hospitalized Surgical Patients: A Cross-Sectional Study at People's Medical College Hospital, Nawabshah

Altaf Hussain Ghumro[1], Inayat Ali Zarzadari[1], Agha Asad Noor[1], Aftab Ahmed Kandhro[2], Shabana Khakheli[3]

[1]*Institute of Microbiology, University of Sindh, Jamshoro, Pakistan*
[2]*Dr. M. A. Kazi Institute of Chemistry University of Sindh, Jamshoro, Pakistan*
[3]*Suleman Roshan College of Nursing and Midwifery, Tandoadam, Pakistan*

ABSTRACT

Background: Pyogenic infections following surgery are important contributors to post-surgical morbidity and may be related to surgical site and biochemical properties of wound exudates. Pus elemental profiling can give information on the wound microenvironment which can direct individual clinical interventions.

Methods: The study design was analytical cross-sectional with 370 consecutive patients (aged 15-80 years) who had infected postoperative wounds in Peoples Medical College Hospital, Nawabshah, Pakistan, between June 2023 and July 2025. Six gastrointestinal surgical sites were aseptically collected pus/wound-exudates. In the wet digestion, sodium, potassium, magnesium, calcium, zinc, copper and chromium were measured in triplicate by inductively coupled plasma optical emission spectrometry. One-way analysis of variance and independent student's *t*-test were used to test the differences by surgical site and age, respectively, and by sex.

Results: The sample comprised 181 males (48.92%) and 189 females (51.08%); 115 patients (31.08%) were aged 61–80 years. Concentrations differed among surgical sites for sodium (p=0.002), magnesium (p=0.001), zinc (p=0.002), copper (p=0.003), and chromium (p=0.004), but not potassium (p=0.104) or calcium (p=0.103). Gut-cancer exudates showed the highest sodium (140.00±7.30 mmol/L), copper (24.00±1.65 µmol/L), and chromium (14.50±1.30 µmol/L). All seven elements were found to be different among the age groups (p=0.001–0.003), all elements except potassium were different by sex.

Conclusion: The work highlights the increased level of sodium, calcium, copper and chromium in the samples of the gut cancer and increased magnesium and zinc in the exudates from the colon and stomach. Significant higher contents of sodium, calcium, copper and chromium are in exudates of females while in males, magnesium and zinc are higher. Sodium, potassium, copper and zinc levels are higher in 61-80 years age group than in other age groups.

Keywords: *Elemental; Exudative; Infection; Metal; Surgical; Wound.*

Received: June 24th, 2026**; Revised:** July 23rd, 2026**; Accepted:** July 29th, 2026

Corresponding Email: ghumro.ah2026@yahoo.com

DOI: https://doi.org/10.59564/amrj/04.03/007

INTRODUCTION

Postoperative pyogenic infections, otherwise known as surgical site infections (surgical site infections are still one of the most common complications after surgery and continue to be a worldwide problem in healthcare. They play a significant role in extended hospital stay, increased use of antimicrobial, and in morbidity and mortality. Recent studies indicate that most surgical site infections are primarily due to endogenous flora of the skin and not contamination from the environment, thus demonstrating the interplay between the host biology and surgical trauma in the mechanism of infection[1,2]. Despite the use of improved aseptic techniques and perioperative care, surgical site infections continue to occur at clinically significant rates, especially in resource poor healthcare environments.

In low and middle income countries, the impact of surgical site infections is greater and is

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

associated with a higher proportion of patients presenting late, poor perioperative infection control practices and poor monitoring infrastructure for the postoperative period[3]. In addition, several risk factors have been demonstrated to affect the incidence of surgical site infection and clinical outcomes, such as surgical site, patient comorbidities, and demographic factors like age and gender[4]. Emergency and abdominal surgery also have been reported to have increased rates of infection, likely because of increased exposure to microbes and contamination of the tissues during surgery, especially gastrointestinal surgery[5]. Existing studies however have been mainly microbiological in nature, and limited exploration of the biochemical environment of infected wounds.

Recent studies have shown a key role for the biochemical composition of wound exudates in addition to microbial load in wound healing and progression of infection. Electrolytes and trace elements such as sodium, potassium, magnesium, calcium, zinc, copper, and chromium are integral to the inflammatory regulation, modulation of oxidant stress, promotion of enzymatic activity, and tissue repair processes[6]. Elemental imbalance of these elements can modify immune reactions and slow wound healing and therefore elemental analysis of pus may be a more powerful tool for understanding wound pathophysiology than is conventional microbiological analysis[7].

Although there is increasing awareness of the role of trace elements in tissue repair, there is still a significant lack of literature that provides a comprehensive elemental profile of the pyogenic wound exudate of postoperative patients. Furthermore, few studies have considered demographic, anatomical and biochemical data together in an analytical model for infection heterogeneity. Thus, in the present study, elemental composition of pus samples from the various surgical sites were evaluated and the association between elemental composition of pus samples with patient's demographic was investigated, which will help to understand the biochemical profile of postoperative wound infections and inform future targeted therapeutic interventions[8,9].

Infection of the postoperative wound continues to be a leading cause of excess morbidity, unnecessary clinical interventions, prolonged hospital stays, delay in recovery, and unnecessary use of antimicrobials, especially in resource limited hospitals. While the identification of the causative organism and the choice of antimicrobial drug is important, microbiological culture alone does not account for all the potential biochemical milieu that is present locally when infection, inflammation, tissue injury, and repair occurs. The composition of postoperative wound exudates may have some elements of the wound microenvironment and can thus offer complementary information. Postoperative wound exudates include inflammatory cells, cellular debris, extracellular fluid, microbial products, electrolytes, and trace elements. Sodium, potassium, magnesium, calcium, zinc, copper, and chromium are involved in fluid and electrolyte balance, cell signaling, enzymatic function, immune function, oxidative stress management, coagulation, collagen formation, angiogenesis and tissue regeneration. The purpose of present study was to compare the elemental composition of pus samples from pyogenic infection following surgery in cases.

MATERIALS AND METHODS

Study Design and Reporting Guideline

This study is an analytical cross-sectional study which was conducted at Peoples Medical College Hospital, Nawabshah, Pakistan during the period of June 2023 to July 2025. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies were used for reporting[10]. Study Population Patients between the ages of 15 and 80 years with clinically infected postoperative wounds after clean or clean-contaminated elective/emergency surgery were the study population. Those that had been treated with a systemic antibiotic prior to the sample collection, who needed reoperation or had not given written informed consent were excluded. Surgical sites included small intestine, colon, duodenum, stomach, spleen and gut-cancer resection sites. Participants were approached sequentially until 370 participants were recruited, with 181 males and 189 females. The sample size and participant recruitment.

Sample Collection and the Recruitment of Participants

A minimum sample size for the estimation of a single population proportion was computed using

$$n = Z^2p\,(1-p)/d^2,$$

Where Z = 1.96 for the 95% confidence level, p = 0.50 (no good estimate of p was available), and d = 0.0523. A size of 352 was initially estimated and to account for the possibility of non-response, 5% was added to this, providing a 370 sample size[11] with 415 patients being assessed during recruitment. Forty-five patients were excluded: 15 had been treated with systemic antibiotics prior to sampling, 11 required reoperation, and 19 refused to participate.

There were 370 participants in the final analysis (Figure 1). A non-probability sampling method, consecutive sampling was employed.

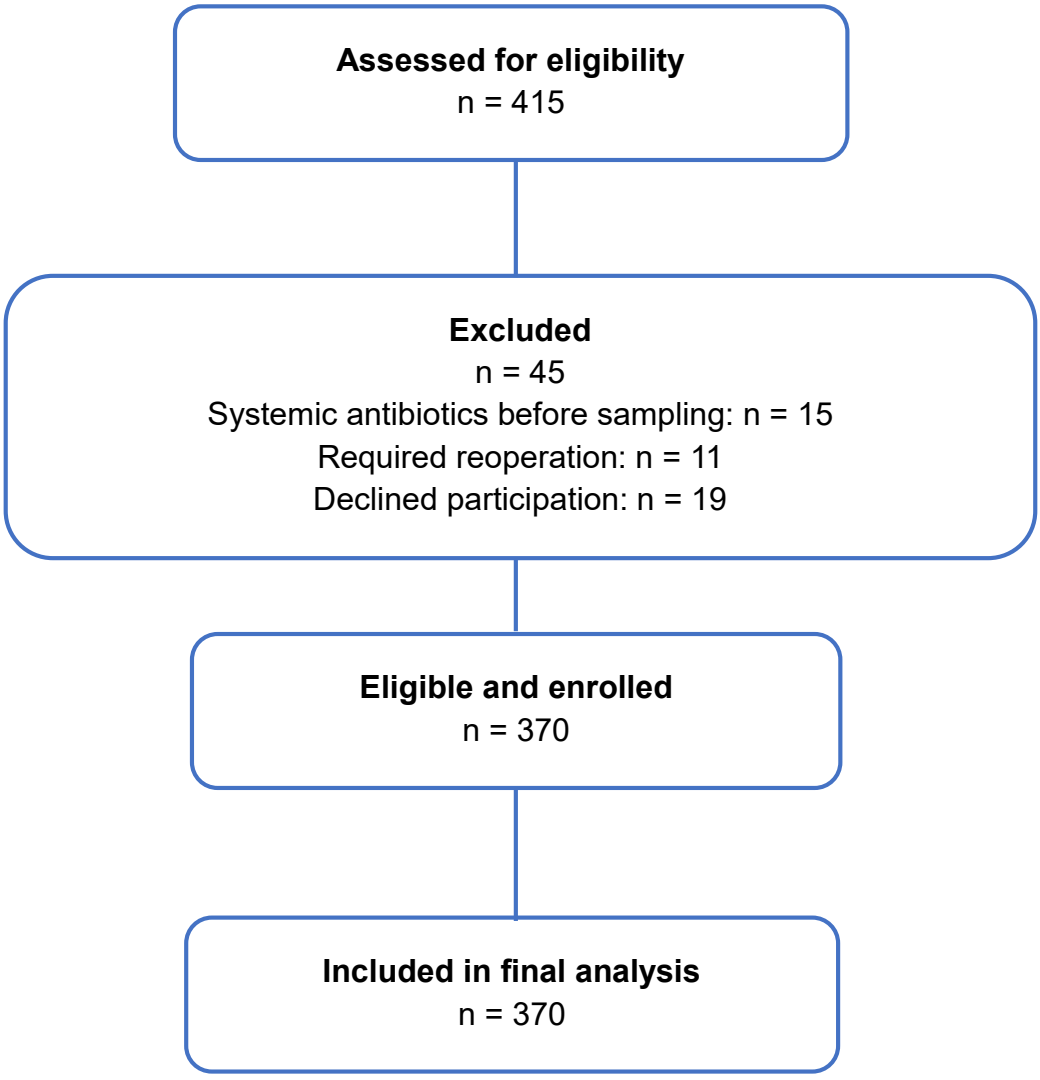

Figure 1. STROBE Flow diagram for selection of participants

Sample Collection

Specimens comprised pus/wound exudate collected directly from infected postoperative wounds, there was no blood or plasma collected for elemental analysis. Pus was aseptically taken with a sterile syringe or sterile swab suitable for the quantity and accessibility of the exudate and processed in accordance with standard biosafety precautions, using

trace-element-free materials. Samples were transferred to the laboratory under 8ºC in vaccination ice box with 15 minutes. The procedures were standardized in order to limit contamination and measurement bias and were followed by all participants. The methods used in this lab were Wet Digestion and ICP-OES Elemental Analysis.

Elemental Analysis

All pus samples were homogenized. Sample preparation was done with some modifications of the method reported by Santos et al. in 2019[12]. About 0.25 g of homogenized pus sample was accurately weighed in a PTFE microwave digestion vessel, and 5 mL concentrated nitric acid (HNO_3, 65%) and 2 mL hydrogen peroxide (H_2O_2, 30%) were added to the vessel. The pre-digestion time was 15-20 minutes at room temperature and the microwave digestion took place for complete oxidation of organic matter, following a controlled temperature program (maximum temperature 180-200°C). The digested solutions were then quantitatively transferred to acid washed polypropylene tubes and diluted to 25 mL using ultrapure deionized water. Reagent blanks and quality control samples were run concurrently and the resulting solutions were analyzed in triplicate by inductively coupled plasma optical emission spectrometry (ICP-OES). Sodium, potassium, magnesium and calcium concentrations were reported in mmol/L; zinc, copper and chromium concentrations were reported in µmol/L.

Data Analysis

Categorical variables were summarized as frequencies and percentages, and elemental concentrations as mean ± standard deviation. Concentrations were compared among surgical sites and age groups using one-way analysis of variance, and among sexes using independent-samples Student t test. The level of significance was determined as $p<0.05$ for 95% confidence interval. All statistical analyses and graphical presentations were done in Microsoft Excel 2013.

RESULTS

In Figure 1 among 370 participants, 181 (48.92%) were male and 189 (51.08%) were female. The 61–80-year group was the largest (n=115, 31.08%), followed by 46–60 years (n=95, 25.68%), 31–45 years (n=90, 24.32%), and 15–30 years (n=70, 18.92%). The distribution of income was split into three groups, lower income (61.62%), middle income (35.68%) and upper income (2.70%) and 57.30% of the population were in the rural part of the country, while 42.70% were in the urban (Figure 2). These percentages reflect the study sample which was infected and are not to be interpreted as incidence or independent risk values as denominators for all operated patients were not available.

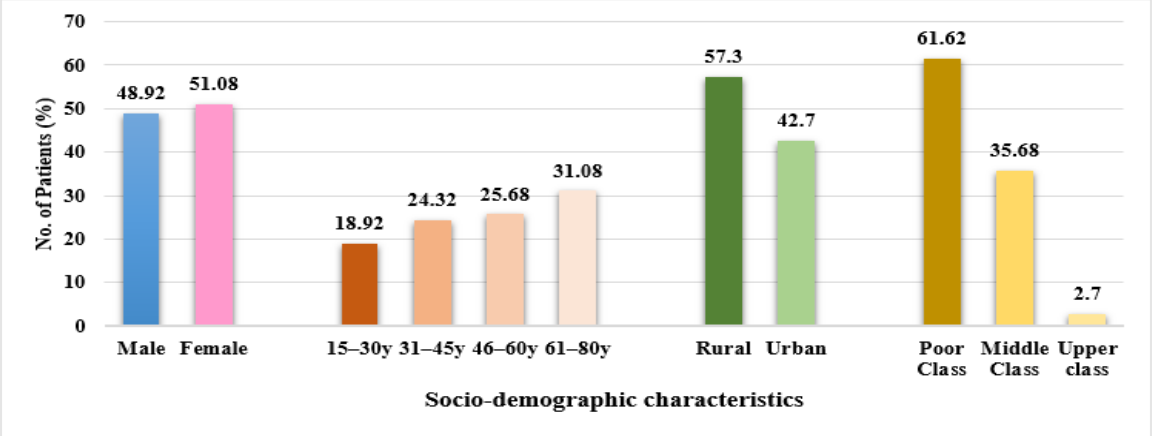

Figure 2. Socio-demographic characteristics of hospitalized patients with postoperative pyogenic infection

There were differences in elemental concentrations between the six surgical sites (Table 1). Significant between-site differences were observed for sodium (p=0.002), magnesium (p=0.001), zinc (p=0.002), copper (p=0.003), and chromium (p=0.004). No significant difference was found between Potassium (p=0.104) and calcium (p=0.103). Gut-cancer exudates had the highest mean sodium (140.00±7.30 mmol/L), magnesium (2.70±0.25 mmol/L), zinc (22.50±1.60 µmol/L), copper (24.00±1.65 µmol/L), and chromium (14.50±1.30 µmol/L).

Table 1. Elemental Profiling of pus samples recovered from Postoperative Pyogenic Infection from different Surgical Sites

Element	Small Intestine n=83 Mean±SD	Colon n = 51 Mean±SD	Duodenum n = 61 Mean±SD	Stomach n = 56 Mean±SD	Spleen n = 59 Mean±SD	Gut Cancer n = 60 Mean±SD	ANOVA p-value (<0.05)
Sodium (mmol/L)	128.50±5.20	135.30±6.80	122.00 ± 7.00	138.50±6.90	125.70 ± 7.10	140.00 ± 7.30	0.002
Potassium (mmol/L)	4.10 ± 0.20	4.55 ± 0.28	3.95 ± 0.27	4.65±0.26	4.00 ± 0.25	4.70 ± 0.28	0.104
Magnesium (mmol/L)	2.10 ± 0.15	2.50 ± 0.22	2.00 ± 0.21	2.60±0.23	2.05 ± 0.20	2.70 ± 0.25	0.001
Calcium (mmol/L)	2.05 ± 0.15	2.40 ± 0.25	2.00 ± 0.22	2.50±0.23	2.10 ± 0.20	2.55 ± 0.25	0.103
Zinc (µmol/L)	19.50 ±1.50	21.50 ± 1.55	18.50±1.50	22.00±1.55	19.00 ± 1.50	22.50 ± 1.60	0.002
Copper (µmol/L)	20.50 ±1.40	23.00 ± 1.60	19.50±1.55	23.50±1.55	21.00 ± 1.50	24.00 ± 1.65	0.003
Chromium (µmol/L)	12.50 ±1.20	14.00 ± 1.25	11. ±1.22	14.20±1.25	12.20 ± 1.20	14.50 ± 1.30	0.004

There was a significant difference (p<0.05) in sodium, magnesium, calcium, zinc, copper and chromium; however, there was no significant difference (p=0.41) between sexes for potassium (Table 2). Females had higher mean sodium (135.50±7.10 mmol/L), calcium (2.40±0.23 mmol/L), copper (23.50±1.65 µmol/L), and chromium (14.50±1.30 µmol/L). Males had higher magnesium (2.95±0.28 mmol/L) and zinc (22.00±1.55 µmol/L).

Table 2. Gender-wise Comparison of Elemental Concentrations in pus samples recovered from Postoperative Pyogenic Infection

Element	Male n =181 Mean ± SD	Female n = 189 Mean ± SD	p-value (<0.05)
Sodium (mmol/L)	130.50 ± 7.20	135.50 ± 7.10	0.028
Potassium (mmol/L)	4.25 ± 0.25	4.30 ± 0.28	0.41
Magnesium (mmol/L)	2.95 ± 0.28	2.70 ± 0.25	0.032
Calcium (mmol/L)	2.25 ± 0.22	2.40 ± 0.23	0.035
Zinc (µmol/L)	22.00 ± 1.55	21.00 ± 1.50	0.030
Copper (µmol/L)	22.00 ± 1.50	23.50 ± 1.65	0.029
Chromium (µmol/L)	13.80 ± 1.25	14.50 ± 1.30	0.031

The differences among the elemental concentrations were significant for all the age groups (p=0.001-0.003) (Table 3). The 61–80-year group had the highest sodium (140.00±7.30 mmol/L), potassium (5.20±0.30 mmol/L), zinc (24.00±1.70 µmol/L), copper (25.00±1.65 µmol/L), and chromium (14.80±1.30 µmol/L), while magnesium and calcium were highest in the 15–30-year group (3.60±0.25 and 2.60±0.25 mmol/L, respectively). These comparisons are unadjusted and attempt to control for the effect of age as an independent variable.

Table 3. Age-wise Comparison of Elemental Concentrations in Pus samples recovered from Postoperative Pyogenic Infection

Element	15–30 years (n=70) Mean ± SD	31–45 years (n=90) Mean ± SD	46–60 years (n=95) Mean ± SD	61–80 years (n=115) Mean ± SD	p-Value (<0.05)
Sodium (mmol/L)	128.50 ± 7.00	135.00 ± 7.50	130.50 ± 7.20	140.00 ± 7.30	0.001
Potassium (mmol/L)	4.15 ± 0.20	4.55 ± 0.25	4.80 ± 0.25	5.20 ± 0.30	0.002
Magnesium (mmol/L)	3.60 ± 0.25	2.50 ± 0.20	3.00 ± 0.22	2.10 ± 0.15	0.001
Calcium (mmol/L)	2.60 ± 0.25	2.40 ± 0.20	2.25 ± 0.20	2.05 ± 0.15	0.003
Zinc (µmol/L)	19.50 ± 1.50	21.00 ± 1.55	22.00 ± 1.55	24.00 ± 1.70	0.001
Copper (µmol/L)	20.50 ± 1.40	22.50 ± 1.50	23.00 ± 1.55	25.00 ± 1.65	0.002
Chromium (µmol/L)	12.50 ± 1.20	13.50 ± 1.25	14.00 ± 1.25	14.80 ± 1.30	0.003

DISCUSSION

Present study aimed to compare the elemental composition of pus samples obtained from pyogenic infections after surgical procedures in cases. The results indicated that the infected subjects were almost equally distributed between females (51.08%) and males (48.92%) as shown in Figure 1. Previous studies showed gender to have little effect on the incidence of surgical site infection.[13] A large German surveillance study found that the adjusted surgical site infection rates were greater for males in general, and for abdominal surgery, and greater for females after some cardiac and vascular surgery. These conflicting findings suggest that sex differences are procedure based and could be greatly influenced by clinical and perioperative rather than sex alone[11].

Older aged groups may be associated with a higher level of surgical activity and exposure, unlike to what was previously reported in regional hospital studies which showed younger and middle aged groups were reported to have postoperative infections.[14] The higher proportion of patients from rural areas may indicate well-known reasons between inadequate access to health care and inadequate postoperative care in rural areas, which has been linked to higher risk of infection in surgical patients[15]. However, in an observational study of elective laparotomy patients, there was no statistically significant relationship between age or sex with the risk of surgical site infectopn, whereas nutritional and perioperative factors were more significant. Thus the present age distribution should not be viewed as suggesting that the older patients were at greater risk of infection as the number of operated patients in each age group was not given, only the number of patients with infection[16].

The lower prevalence among urban residents on the other hand, may suggest better perioperative care and hygiene practice[17]. There was a socioeconomic difference, with higher infection rates in lower income patients. This trend may be due to poor nutrition status, lack of knowledge about postoperative care and inability to seek timely medical attention, all of which are recognized risk factors for surgical site infections[18,19]. Minimal involvement of the

upper class highlights the importance of socioeconomic status for health outcomes after surgery, consistent with what other epidemiological studies have found[20]. Thus, the rural population and low income should be considered as population characteristics of the infected group and not as separate risk factors. Delayed presentation, transportation problems, nutrition, postoperative follow-up, and access to wound care are potential factors which are still plausible and should be investigated analytically[21].

The large differences in elemental concentrations in various surgical locations (in Table 1) indicate that the biochemistry of the wounds may be affected by both the tissue type and the surgical site. The significant high amounts of sodium, copper and chromium in the cancer of the gut and stomach wounds may be related to the high level of inflammation and metabolism in these tissues and the literature suggests that these elements have a crucial role in tissue repair, enzymatic activity and immune response[22,23]. These sites had relatively low concentrations of magnesium and zinc, and relatively high concentrations in colon and stomach wounds, which is consistent with previous studies indicating that these minerals play an important role in collagen synthesis and oxidative stress control during wound healing[24,25]. The uniformity observed in the site of calcium and potassium has suggested a regulation of potassium during the healing and infection process, which is consistent with the reported regulation of electrolyte in postoperative exudates[26].

The concept of a local biological environment varies between surgical sites, with recent genomic data demonstrating that endogenous microbes from the patient play a role in postoperative infections observed at each site. The present study does not provide direct support, however, the study investigated the origin of microorganisms rather than concentrations[27]. Cancer, malnutrition, contamination level and complexity of the operation have been identified as independent factors associated with surgical site infection following abdominal surgery, suggesting that the gut-cancer group may have differences in clinical characteristics from those undergoing non-cancer surgery. Therefore, it is important that the higher concentrations of the elements were not associated with the malignant tissue or anatomical site alone, but should be considered in the context of these factors[28]. In patients with burn, a significant amount of Zn and Cu have been found to be excreted in the wound exudate and there is also a reduction in the systemic Zn and Cu status. This is significant because, a high level of an element in wound fluid does not necessarily mean that the systemic nutritional status is adequate or that there is any redistribution or loss due to wound leakage. Therefore, the present findings are generalizable to the evidence that the exudate from a wound can contain clinically relevant amounts of trace elements but serum levels and nutritional status would be needed to ascertain whether the patients were deficient[29].

The higher concentrations of sodium, calcium, copper and chromium in females (Table 2) may be due to a gender difference in metabolism and/or inflammatory responses to the wounds. Whether the role of these elements in enzymatic function, modulation of oxidative stress and tissue repair is different between the sexes has been reported, and if so, it is one that may be specific to these functions[30,31]. On the other hand, magnesium and zinc were raised in males, which aligns with their known functions in collagen formation, immune regulation, and antioxidant protection in wound healing process[32,33]. There was no significant difference in potassium, which agrees with the results of other studies in which potassium was closely regulated in both sexes[34].

There are also conflicting results with regard to sex in the broader surgical site infection literature; some large surveillance studies have observed higher rates of surgical site infection in males following abdominal surgeries, while more comprehensive and detailed studies have controlled for obesity, malnutrition, cancer, operative time, contamination, and transfusion and reported no independent association with sex. This may therefore be due to the different distributions of age, type of surgery, malignancy, lactation status, nutritional status, renal function, intravenous fluid and comorbidities between male and female. They should be interpreted as a difference, rather than as evidence that biological sex directly affected the composition of wound-exudates[28].

The great age-dependent variation in elemental concentrations indicates that there may be physiological age changes affecting the biochemistry of wounds (Table 3). Increased levels of sodium and potassium could represent changes in fluid balance and cell homeostasis that occur during tissue repair as previously seen in recovery after surgery[35]. Higher levels of Cu and Zn in older groups might suggest greater demand for enzymes and/or oxidative stress in the course of chronic or complicated healing[34]. The levels of magnesium and

calcium vary with age in middle aged patients, and is related to their function in enzymatic activity, collagen formation and inflammatory regulation, which varies with metabolic rates and nutrition status[36,37]. The age-related differences seen in the observed responses emphasize the need to take patient's age into consideration when assessing the potential for wound healing and when planning postoperative care, such as the use of trace elements as supplements[35,36]. Older age has been previously linked to this more directly, including obesity, malnutrition, malignancy, contamination, operative time, laparotomy, and blood transfusion. Data on the elemental findings related to age should, therefore, be treated as exploratory and should be confirmed by multivariable regression analysis if possible using age as a continuous variable[26]. Wound exudate elemental concentrations may be influenced by cellular breakdown, inflammatory cells, blood contamination, bacterial metabolism, local tissue composition, wound irrigation, dressings, and fluid movement.

Because of its cross-sectional design and its monocentric nature, it is not possible to determine the cause of the changes and the findings may not be generalizable. There are no non-infected postoperative-wound controls available, therefore a normal exudate elemental profile cannot be defined. However, serum or tissue elemental levels, nutritional status, renal function, and exposure to intravenous fluids, comorbidities, wound contamination class, and microbiological findings were not measured and therefore local exudate concentrations cannot be interpreted as systemic deficiency or associated to specific pathogens. The analyses were unadjusted; differences by surgical site, sex or age might be due to confounding. Control groups, as well as standardization of laboratory protocols, microbiological characterization and multivariable modelling should be incorporated in future multicenter studies.

CONCLUSION

Significant differences across surgical site were seen for Na, Mg, Zn, Cu and Cr, but not for K and Ca. Additionally, the differences between the sexes were determined, with the females having more sodium, calcium, copper, and chromium and the males having more magnesium and zinc. These results reflect biochemical variability of exudates from infected postoperative wounds, but do not prove causality or clinical value. Patient management should be guided by elemental profiling which needs to be validated in controlled, multicenter studies.

ACKNOWLEDGMENTS

The authors would like to thank the staff of the Institute of Microbiology, surgical staff and surgeons of Peoples Medical College Hospital, Nawabshah and their patients for cooperation in the collection of pus samples. This study was greatly aided by the assistance of the trained staff.

Author Contributions

AHG: Conceived and designed the study, supervised the research, and contributed to manuscript drafting.
IAZ: Assisted in study design, data collection, and data analysis.
AAN: Contributed to the literature review, data interpretation, and manuscript preparation.
AAK: Assisted in data collection, data entry, and manuscript editing.
SK: Critically revised the manuscript, provided overall guidance, and approved the final version.

Funding Source

No sources.

Conflict of Interest

The authors declared no conflict of interest.

Data Availability

Data will be available on request.

Ethical Approval

The protocol was approved by the Institute of Microbiology University of Sindh, Jamshoro (Ref No. IMB/ 217 /2023).

REFERENCES

1. Leandri M, Vallicelli C, Santandrea G, et al. Postoperative infections after appendectomy for acute appendicitis: the surgeon's checklist. Antibiotics (Basel). 2025;14(9):954. https://doi.org/10.3390/antibiotics14090954
2. Wenzel RP. Surgical site infections and the microbiome: an updated perspective. Infect Control Hosp Epidemiol. 2019;40(5):590-596. https://doi.org/10.1017/ice.2018.363
3. Graf PD, et al. Bacterial reservoir in deeper skin is a potential source for surgical site and biomaterial-associated infections. J Hosp Infect. 2023;140:62-71. https://doi.org/10.1016/j.jhin.2023.07.014
4. Khan FU, Fang Y, Khan Z, et al. Occurrence, associated risk factors, and treatment of surgical site infections in Pakistan: a hospital-based study. Eur J Inflamm. 2020;18:2058739220960547. https://doi.org/10.1177/2058739220960547
5. Gupta P, Lal P, Manchanda V. Microbial profile from peritoneal fluid and surgical wounds in perforation peritonitis. Postgrad Med J. 2023;99(1171):1-8. https://doi.org/10.1177/00494755221150537
6. Misha G, Chelkeba L, Melaku TB. Bacterial profile and antimicrobial susceptibility patterns of isolates among patients diagnosed with surgical site infection at a tertiary teaching hospital in Ethiopia: a prospective cohort study. Ann Clin Microbiol Antimicrob.

2021;20:33. https://doi.org/10.1186/s12941-021-00440-z
7. Linju M, Rekha M. Role of inorganic ions in wound healing: an insight into localized delivery approaches. Ther Deliv. 2023;14(10):649-667. https://doi.org/10.4155/tde-2023-0036
8. Majid E, Pathan S, Zuberi BF, et al. Comparison of honey and povidone iodine dressings in post-cesarean surgical site wound infection healing. Pak J Med Sci. 2023;39(6):1803-1808. https://doi.org/10.12669/pjms.39.6.7499
9. Arija D, Mirani SH, Abideen ZU, et al. Serum C-reactive protein, procalcitonin, and lactate levels as predictors of postoperative surgical site infections. Pak J Med Health Sci. 2026;20(3):26-32. https://doi.org/10.53350/pjmhs02026203.5
10. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. https://doi.org/10.1371/journal.pmed.0040296
11. Lwanga SK, Lemeshow S. Sample Size Determination in Health Studies: A Practical Manual. Geneva: World Health Organization; 1991.
12. Santos HM, Caires ARL, Lemos VA, et al. Microwave-assisted digestion using diluted HNO_3 and H_2O_2 for multi-element determination in guarana samples by ICP OES. Food Chem. 2019;275:457-463. https://doi.org/10.1016/j.foodchem.2018.09.114
13. Aghdassi SJS, Schröder C, Gastmeier P. Gender-related risk factors for surgical site infections: results from ten years of surveillance in Germany. Antimicrob Resist Infect Control. 2019;8:95. https://doi.org/10.1186/s13756-019-0547-x
14. Alemayehu MA, Azene AG, Mihretie KM. Time to development of surgical site infection and its predictors among general surgery patients: a prospective follow-up study. BMC Infect Dis. 2023;23:334. https://doi.org/10.1186/s12879-023-08301-0
15. Mengistu DA, Alemu A, Abdukadir AA, et al. Global incidence of surgical site infection among patients: systematic review and meta-analysis. Am J Med Sci. 2023;83:104324. https://doi.org/10.1177/00469580231162549
16. Huda F, Shasheendran S, Basu S, et al. Risk factors of surgical site infection in elective laparotomy in a tertiary care center: an observational study. Int J Burns Trauma. 2022;12(3):106-113.
17. PakSurg1 investigators. Determining the epidemiology and risk factors of surgical site infections in Pakistan: a multicentre prospective cohort study. BMJ Open. 2023;13(7):e070831. https://doi.org/10.1136/bmjopen-2022-070831
18. Rezaei AR, Zienkiewicz D, Rezaei AR. Surgical site infections: a comprehensive review. J Trauma Inj. 2025;38(2):71-81. https://doi.org/10.20408/jti.2025.0019
19. Ansari S, Hassan M, Barry HD, et al. Risk factors associated with surgical site infections: a retrospective report from a developing country. Cureus. 2019;11(6):e4801. https://doi.org/10.7759/cureus.4801
20. Tahir MI, Waheed MK, Nadeem S, et al. Assessment of surgical site infection rates in emergency versus elective procedures: a cross-sectional study. Ann Pak Med Dent Coll. 2024;1(3):14-19. https://doi.org/10.53350/Annalspakmed.1.3.4
21. GlobalSurg Collaborative. Surgical site infection after gastrointestinal surgery in high-income, middle-income, and low-income countries: a prospective international multicentre cohort study. Lancet Infect Dis. 2018;18(5):516-525. https://doi.org/10.1016/S1473-3099(18)30101-4
22. Mai K, Maverakis E, Li J, Zhao M. Maintaining and restoring gradients of ions in the epidermis: the role of ion and water channels in acute cutaneous wound healing. Adv Wound Care (New Rochelle). 2023;12(12):780-792. https://doi.org/10.1089/wound.2022.0128
23. Hou HT, Xue LG, Zhou JY, et al. Alteration of plasma trace elements during and after coronary artery bypass grafting surgery. J Trace Elem Med Biol. 2020;62:126612. https://doi.org/10.1016/j.jtemb.2020.126612
24. Dixit R, Chaudhary NK, Mishra PK, et al. Study on blood serum levels of heavy and trace metals in chronic non-healing wounds. Integr Cancer Ther. 2022;23:15347346221074161. https://doi.org/10.1177/15347346221074161
25. Dalisson B, Barralet J. Bioinorganics and wound healing. Adv Healthc Mater. 2019;8(18):1900764. https://doi.org/10.1002/adhm.201900764
26. Mukherjee K, Kavalukas SL, Barbul A. Nutritional aspects of gastrointestinal wound healing. Adv Wound Care (New Rochelle). 2016;5(11):507-515. https://doi.org/10.1089/wound.2015.0671
27. Long DR, Bryson-Cahn C, Waalkes A, et al. Contribution of the patient microbiome to surgical site infection and antibiotic prophylaxis failure in spine surgery. Sci Transl Med. 2024;16(742):eadk8222. https://doi.org/10.1126/scitranslmed.adk8222
28. Zwicky SN, Gloor S, Tschan F, et al. No impact of sex on surgical site infections in abdominal surgery: a multi-center study. Langenbecks Arch Surg. 2022;407(8):3763-3769. https://doi.org/10.1007/s00423-022-02691-6
29. Voruganti VS, Klein GL, Lu HX, et al. Impaired zinc and copper status in children with burn injuries: need to reassess nutritional requirements. Burns. 2005;31(6):711-716. https://doi.org/10.1016/j.burns.2005.04.026
30. Lansdown ABG. Sequential changes in trace metal, metallothionein, and calmodulin concentrations in healing skin wounds. J Anat. 1999;195(3):375-386. https://doi.org/10.1046/j.1469-7580.1999.19530375.x
31. Sirkka T, Skiba JB, Apell SP. Wound pH depends on actual wound size. arXiv. 2016. https://arxiv.org/abs/1601.06365
32. Linju MC, Rekha MR. Role of inorganic ions in wound healing: insights into localized delivery approaches. Ther Deliv. 2023;14(10):649-667. https://doi.org/10.4155/tde-2023-0036
33. Adjepong D, Jahangir S, Malik BH. The effect of zinc on post-neurosurgical wound healing: a review. Cureus. 2020;2(1):e6770. https://doi.org/10.7759/cureus.6770
34. Salvo J, Sandoval C. Role of copper nanoparticles in wound healing for chronic wounds: literature review. Burns Trauma. 2022;10:tkab047. https://doi.org/10.1093/burnst/tkab047
35. Tabriz N, Fried D, Uslar V, Weyhe D. Impact of preoperative calcium and magnesium supplementation on quality of life and hypocalcemia post-thyroidectomy. Endocrinol Diabetes Metab. 2026;9(1):e70129. https://doi.org/10.1002/edm2.70129
36. Ranire-Maguire M, Caprioli R, Taylor C, Rodger R, Thomas A, Oropallo A. Exudate management with a next-generation multilayered foam dressing in traumatic and surgical wounds: a prospective multicentre study. J Wound Care. 2026;35:392-398. https://doi.org/10.12968/jowc.2026.0048
37. Palmieri B, Vadalà M, Laurino C. Nutrition in wound healing: investigation of the molecular mechanisms, a narrative review. J Wound Care. 2019;28(10):683-693. https://doi.org/10.12968/jowc.2019.28.10.683

Clinical Utility

Allied Medical Research Journal

Vol:4 No.3: 65-72

© The Author (s) 2026

Hybrid Ensemble Learning for Multi-Class Chest X-Ray Classification Using Deep CNN

Abdul Rehman Khan Tareen[1], Muhammad Laiq Ur Rahman Shahid[1], Furqan Shaukat[1], Muhammad Hamza Zafar[2], Soban Abu Khifs[3]

[1]*Dept of Elecrical Engineering, University of Engineering and Technology Taxila, Pakistan*
[2]*Dept of Engineering Sciences, University of Agder, Grimstad, Norway*
[3]*Dept of Medical Surgery, Abbas Institute of Medical Sciences, Muzaffarabad, AJK, Pakistan*

ABSTRACT

Background: COVID-19, pneumonia, and TB (tuberculosis) are still the big killers of people suffering from chest disease and continue to be a serious health challenge globally. A timely diagnosis leads to timely treatment and improved patient outcomes. Chest X-ray (CXR) imaging is widely used for diagnostic purposes due to its speed, low cost, and availability in most healthcare facilities. Manual reading of CXR images is, however, challenging because of the similarity in the presentation of radiographic features across various chest diseases.

Methods: This research introduces a hybrid ensemble learning approach to classify chest X-ray images into four classes—Normal, COVID-19, Pneumonia, and Tuberculosis. Three Deep CNN network models, namely Xception, AlexNet, and EfficientNet-B0, were used for deep feature extraction. Additionally, texture features of the images were extracted using Gabor filters. The deep and texture features were combined and classified using logistic regression and a stacking ensemble learning approach. A publicly available chest X-ray image database containing 7,135 X-rays was used, with six-fold stratified cross-validation to assess the proposed approach.

Results: The ensemble models outperformed the individual CNN models. The Average Ensemble produced the best results with an accuracy of 91.18%, an Average Precision (AP) of 96.75%, and an Area Under the ROC Curve (AUC) of 98.87%. The proposed model performs well across all four disease classes. It exhibited high sensitivity in detecting Tuberculosis with considerable stability in classifying Normal, COVID-19, and Pneumonia.

Conclusion: The proposed framework demonstrates the effectiveness of integrating deep learning features, Gabor texture features, and ensemble learning for improved chest X-ray image classification. This can help computer-assisted diagnostics systems and aid medical workers in identifying chest diseases early.

Keywords: Chest X-ray, Deep Learning, Convolutional Neural Networks, Ensemble Learning, Medical Image Classification, Tuberculosis, Pneumonia.

Received: June 25th, 2026**; Revised:** July 23rd, 2026**; Accepted:** July 28th, 2026

Corresponding Email: khan0229508@gmail.com

DOI: https://doi.org/10.59564/amrj/04.03/000

INTRODUCTION

Chest diseases, including pneumonia, COVID-19, and tuberculosis (TB), remain among the leading causes of illness and death worldwide. The burden of these diseases is particularly high in developing countries, where access to advanced healthcare facilities may be limited. Early and accurate diagnosis is crucial for effective treatment and disease control. Chest X-ray (CXR) imaging is one of the most commonly used diagnostic tools for chest disease assessment because of its low cost, rapid acquisition, and widespread availability in clinical practice.

The medical field has been greatly helped by many artificial intelligence (AI) technologies for medical image interpretation. One type of deep learning network widely used in chest disease recognition is a convolutional neural network, as it can learn discriminative features from CXRs[2-5]. A large body of research has employed deep learning & ensemble learning principles for tuberculosis detection, achieving strong classification performance[1-4].

Transfer learning streamlines convolutional neural network training by leveraging disease-specific features from large-scale medical datasets, such as Xception and EfficientNet[5-10]. Furthermore, there have been attempts to improve the model's performance in terms of accuracy and generalization across various projects in the context of image improvement and explainable artificial intelligence[6,11,12].

Recent research indicates that using multiple deep learning models to classify chest X-rays can improve reliability. The purpose of ensemble learning is to overcome the drawbacks of single learning models by combining model predictions, thereby achieving better generalization and classification performance. Besides, hybrid models that combine deep learning with other feature extraction methods have also been shown to enhance the detection of chest diseases, particularly those with similar radiographic features[25,26].

However, a few studies on this topic continue to focus on classifying just two classes, most prominently the

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0), which permits others to share, copy, redistribute, and adapt the work for non-commercial purposes, provided the original author(s) and source are credited appropriately. Further details are available on the official AMRJ Open Access policy page: https://ojs.amrj.net/index.php/1/14.

distinction between healthy and TB patients' CXRs[1-8]. On the other hand, several methods in the literature rely on a single CNN without considering the advantages of using hand-made textures for the classification problem. Since each of the three diseases, COVID-19, Pneumonia, and TB, has different radiological characteristics, deep + texture features can increase the ability to recognize the disease.

In this paper, a framework for classifying 4 types of chest diseases is proposed that combines texture and deep features using deep learning and ensemble learning techniques. Deep features such as Xception, AlexNet, and Efficient Net will be processed with Gabor filtering, and classification will be performed using logistic regression and a stacked ensemble for four classes (Normal, COVID-19, Pneumonia, and Tuberculosis).

Building on the above innovations, this research presents a hybrid model for automated chest disease classification that integrates transfer learning, Gabor texture analysis, feature fusion, and ensemble learning. The proposed method combines discriminative, complementary information from a variety of pre-trained CNN classifiers with hand-crafted texture features to improve classification performance. The framework was developed for computer-aided diagnosis and also helps health workers in promoting early detection of chest diseases[25,28].

CNN models are commonly used for chest X-ray classification, and some studies have applied ensemble learning methods for the same task; however, most studies have focused on binary classification or on a specific deep learning model. Some approaches use a combination of CNN models without paying much attention to texture information, thereby enhancing the representation of chest X-ray images. To achieve this, a hybrid model is proposed in this study that merges three pre-trained CNN deep features with Gabor deep texture features through deep feature fusion and ensemble. The proposed approach differs from many others in the literature in a few respects: it incorporates a single set of design rules for the classification problem of Normal, COVID-19, Pneumonia, and Tuberculosis. With this combination, the model obtains complementary deep and texture information, achieving good classification performance while preserving its simplicity and applicability.

Below are some chest images from 4 classes in this experiment. The Normal class represents clear lung images, and the COVID-19 and Pneumonia classes represent pulmonary opacities or infiltrates. However, tuberculosis pictures consist of localized and cavity appearance.

The key points about the contributions are highlighted below:

- Explores the possibility of a hybrid system in which Deep CNN features, Gabor texture features, feature fusion, and ensemble learning are incorporated into a four-class chest X-ray classification. The proposed system differs from previous studies, which focused on a few classes or used a single CNN model, by supporting multi-class disease classification within a single system.
- For feature extraction, we use deep learning models; in this case, we use Xception, AlexNet, and EfficientNet. For the texture feature, we use a Gabor filter. By using a combined algorithm that fuses deep and texture features, the classifier's discriminative ability can be enhanced.
- A total of 7,135 chest X-ray images are used to test the performance of the proposed framework. Evaluated metrics are Accuracy, Precision, Recall, F1-score, ROC-AUC, and Precision-Recall curves.

METHODOLOGY

The proposed framework is intended to categorize chest X-ray images into the Normal, COVID-19, Pneumonia, and Tuberculosis classes. It employs deep feature extraction using pre-trained CNN models, Gabor filter-based texture feature extraction, and stacked ensemble learning techniques.

The process flow, which indicates the steps of the proposed system, is shown in Figure 2 below. Preprocessed images will undergo two types of feature extraction techniques. Pre-trained CNN models will be used for deep feature extraction, and a Gabor filter will be used for texture feature extraction. Then, a Logistic regression Classifier will be trained using the obtained feature vectors. Finally, a stacked classifier based on ensemble learning will be used to obtain the classification result.

All chest X-ray images were preprocessed to ensure uniformity of the input to deep learning models before feature extraction. First, all images were resized to 224×224, the input size required by the selected CNN models. Normalization of Pixel values for improving training stability. Several data augmentation methods, such as random rotations, translations, and flips, were applied only to images in the training set to diversify the data and limit overfitting. This involved random rotation, horizontal reflections, magnification, and horizontal and vertical shifts. No efforts were made to enhance the testing images to provide a fair test of the proposed framework.

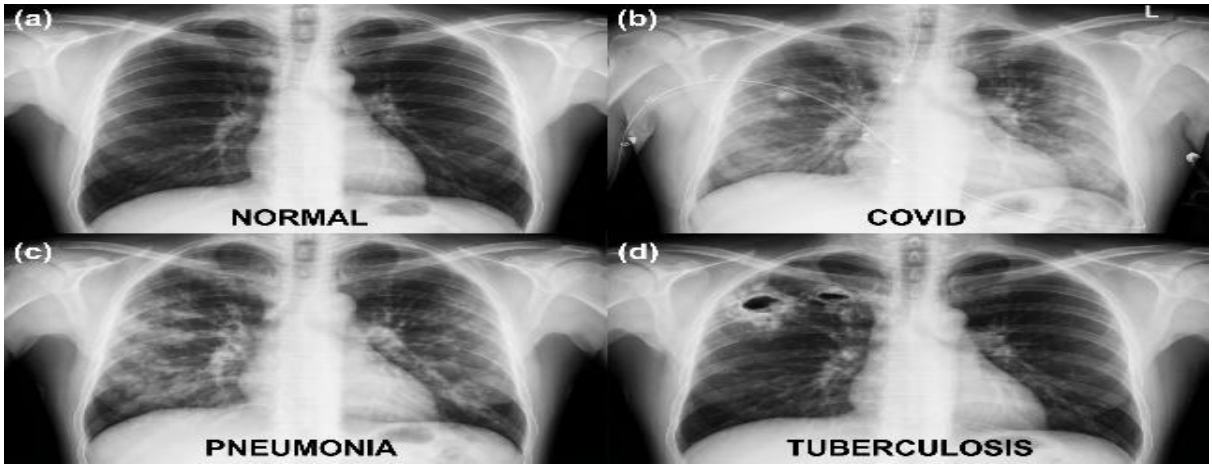

Figure 1: Representative chest X-ray images used in this study

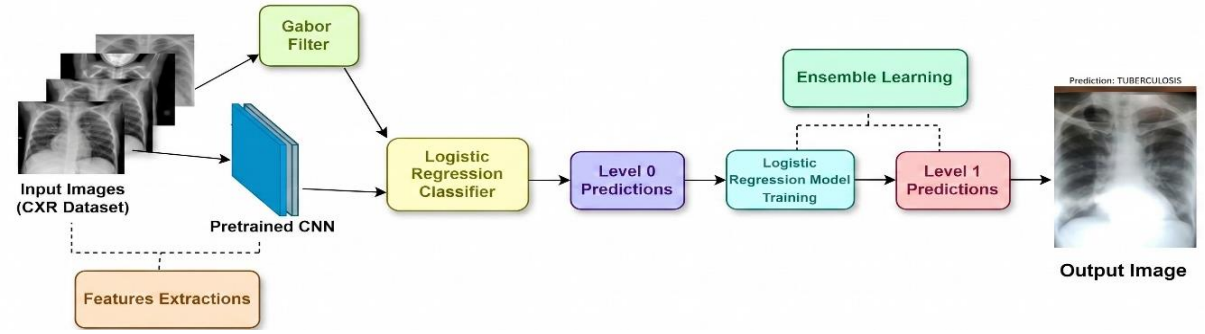

Figure 2: Block Diagram of the proposed method

The training set is augmented to improve the model's generalizability. The data augmentation techniques used are: Random rotation, Horizontal flipping, Zooming, Width shifting, and Height shifting. Those processes are applied to the training images to achieve a critical mass of diversity without losing the signal in the training data.

All deep learning models were initialized from an ImageNet-pretrained model and fine-tuned on the chest X-ray dataset. All images were scaled to 224x224 during training. The models were trained with the Adam optimizer at a learning rate of 0.0001 and a batch size of 32, using categorical cross-entropy loss and a Softmax activation for multi-class classification. They underwent training for 100 epochs, with the model achieving the highest validation score; this model was selected for further evaluation.

The models chosen were Xception, AlexNet, and EfficientNet-B0, each with a different network design and trained to learn different types of image features. Efficiently capturing complex patterns in chest X-ray images, Xception is effective at understanding the deep spatial features through depth wise separable convolutions. The AlexNet architecture is simple and well-known for efficiently learning low-level image features. Compound scaling of network depth, width, and resolution yields EfficientNet-B0, which offers a good trade-off between classification performance and computational cost. By combining three models, the proposed model learns complementary features from different network architectures, thereby enhancing the overall feature representation before feature fusion and ensemble learning.

For the three CNN models, different training approaches were used to improve training. Transfer learning was used to initialize Xception and EfficientNet-B0 with weights from a large-scale pre-trained model (ImageNet). In the first phase, the additional classification layers were trained, while the backbone networks were frozen. During fine-tuning, only the last 30 layers of both Xception and EfficientNet-B0 were unfrozen and updated, while the other layers, such as the Batch Normalization layers, were kept frozen. In contrast, AlexNet was trained from scratch, and all 10 trainable layers (5 convolutional layers, 2 Batch Normalization layers, and 3 fully connected layers) were optimized during training. This training procedure enabled the pre-trained models to fine-tune on the chest X-ray image dataset and to learn task-specific image features without relying on the existing knowledge or experience of AlexNet.

For each disease, class weights have been provided to ensure it has the same proportion of instances as the other classes in the training set. The heart and soul of the proposed framework is feature extraction. Both deep learning and texture-based feature extraction techniques are employed to provide a comprehensive feature set for chest X-rays.

Although deep learning models perform very well at learning semantic features of an image, texture characteristics are also of great importance, particularly in medical image applications such as the detection of pulmonary diseases.
In this study, texture features are learned from chest X-ray images using a two-dimensional Gabor filter. The mathematical equation of the Gabor filter can be expressed as:

$$G(x,y) = exp\left(-\frac{x'^2 + \Upsilon^2 y'^2}{2\sigma^2}\right) cos\left(2\pi\frac{x'}{\lambda} + \emptyset\right)$$

Where λ represents wavelength, $\emptyset$ ε represents orientation, π represents phase offset, σ represents Gaussian spread, and Υ represents spatial aspect ratio.

The final feature representation will be composed of CNN and Gabor features:

$$F_{Final} = [F_X, F_A, F_E, F_G]$$

Where F_G represents the Gabor texture features.
The features from the Xception, AlexNet, and EfficientNet-B0 models were stacked with Gabor texture features; this stacking was referred to as feature concatenation. This involves combining the feature vectors to create a representation of the whole without modifying the constituent feature values. The feature vector is a fusion of the deep learning features and the texture features. The fused features were then fed into the Logistic Regression classifier for disease diagnosis.

The feature vector extracted from each model had a different size. Before going into classification, the Global Average Pooling layer yielded a 2048-dimensional feature vector, and AlexNet yielded a 1024-dimensional feature vector from its last fully connected layer. Global Average Pooling outputs a 1280-dimensional feature vector from the input of EfficientNet-B0. The Gabor feature extraction stage generated a 48-dimensional vector of texture features based on statistical values of filter responses. These were all combined into a fused 4400-dimensional feature vector, which was used as input to the Logistic Regression classifier.

The set of features is combined into a single vector and used as input to a Logistic Regression classifier that produces class-label probabilities (i.e., the probability of chest disease).

$$L_o = [P_1, P_2, P_3, P_4]$$

Where P_1, P_2, P_3, and P_4 represent the probabilities of normal, Covid, Pneumonia, and Tuberculosis, respectively.

In this study, two ensemble learning strategies were tested: The Stacking Logistic Regression Ensemble and the Average Ensemble. For the stacking approach, all the prediction probabilities produced by the individual CNN models were considered as input to a Logistic Regression

meta-classifier to learn how to combine predictions from the base learners. For prediction in the Average Ensemble, each CNN model's predicted probability was averaged, and the class with the highest average probability was selected. The performance of multiple classifiers in combination was evaluated to determine whether combining them would yield a more robust and overall better classification.

The Kaggle Chest X-ray dataset, a publicly available dataset, was used for all experiments and comprises chest X-ray images from four classes: Normal, COVID-19, Pneumonia, and Tuberculosis. In total, there are 7,135 images. For this, 1,341 Normal, 460 COVID-19, 3,875 Pneumonia, and 650 Tuberculosis images were used to train the model. Testing Set includes 234 images of Normal persons, 106 images of COVID-19 persons, 390 images of Pneumonia persons, and 41 images of TB persons. Furthermore, a validation set was used to develop the model, comprising 8 Normal, 10 COVID-19, 8 Pneumonia, and 12 Tuberculosis images. Class weights were used during training to mitigate the effects of class imbalance, as the Pneumonia class is much larger than the others.

TABLE 1: Distribution of the chest X-ray dataset used for training, validation, and testing.

Class	Training	Testing	Validation	Total
Normal	1341	234	8	1583
COVID-19	460	106	10	576
Pneumonia	3875	390	8	4273
Tuberculosis	650	41	12	703
Total	6320	771	38	7135

The distribution of images across the four classes is imbalanced, with the largest number of images in the Pneumonia class. For these reasons, when all deep learning models were trained, class weights were applied to reduce this imbalance during training. This enables the model to be successfully trained by both majority and minority classes. To make the training data more suitable for the model training and validation, a 6-fold stratified cross-validation was used to ensure the data in each fold had a similar class distribution.

This study does not include any patient identifiers in the publicly available Kaggle chest X-ray dataset. Hence, it was not possible to confirm patient-level separation between the training and validation sets. The experiments were conducted according to the specified data set structure given with the data set. Sixfold stratified cross-validation was used to train the models, creating folds with an equal distribution of the four disease classes, thereby minimizing the risk of biased performance estimates.

RESULTS

The results of all individual and ensemble CNN models were evaluated using accuracy, sensitivity, specificity, AP, and AUC. All results from individual and ensemble models are summarized in Table 2 below.

TABLE 2: Performance comparison of the individual CNN models and ensemble methods.

Model	Accuracy (%)	AP (%)	AUC (%)
Xception	90.27	95.29	98.22
AlexNet	90.66	95.71	98.5
EfficientNet-B0	83.53	92.01	96.87
Stacking Logistic Regression Ensemble	91.05	95.88	98.19
Average Ensemble	91.18	96.75	98.87

Among individual CNN models, AlexNet achieved the highest accuracy, 90.66%, followed by Xception at 90.27%. As for the Ensemble models, the best accuracy was observed with the Average Ensemble Model at 91.18%, followed by the Stacking Logistic Regression Ensembles at 91.05%.

Based on the results, the ensembles have outperformed the CNNs. In particular, the Average Ensemble exhibited the best performance with an AP of 0.9675 and an AUC of 0.9887, thereby highlighting its outstanding classification capability.

Compared with CNN models, AlexNet performed best, achieving 90.66% accuracy, followed by Xception at 90.27%. Both models achieved a high level of accuracy (100%) when it came to diagnosing Tuberculosis (TB) cases. The EfficientNet-B0 model achieved the lowest accuracy, at 83.53%, but performed excellently in classifying COVID-19 cases.

The Average Ensemble, in terms of performance, had the highest accuracy of 91.18%, sensitivity of 93.20%, and specificity of 95.84%. Similarly, the Stacking Logistic Regression Ensemble achieved an accuracy of 91.05% and performed well across all diseases, particularly in the diagnosis of COVID-19 and TB.

To evaluate the individual effect of each component, an ablation study was conducted, in which different configurations of the suggested framework were investigated. Four experimental conditions were tested, including the effects of Gabor features, individual CNN models, feature fusion, and ensemble learning. The performance of each configuration with the identical evaluation settings is summarized in Table 3.

TABLE 3: Ablation study showing the contribution of Gabor features, feature fusion, and ensemble learning to the proposed framework.

Model	Accuracy	Precision	Recall	F1-Score
Gabour Feature Only	0.7030	07103	0.7263	0.7038
EfficientNet-B0	0.8444	0.8616	0.8467	0.8489
AlexNet	0.8988	0.8872	0.9132	0.8950
Xception	0.9040	0.9120	0.9213	0.9130
CNN Stacking	0.9053	0.9174	0.9215	0.9151
Feature Fusion + Gabor + CNN	0.8923	0.8740	0.9032	0.8800
Average Ensemble (Proposed)	**0.9118**	**0.9301**	**0.92286**	**0.9263**

Results reveal that every element of the proposed framework contributed to the overall performance. The classification accuracy of the Gabor features was the lowest, indicating that the texture data alone could not

enable effective disease classification. Of all the CNN models, Xception had the highest scores, followed by AlexNet and EfficientNet-B0. The classification results showed that the ensemble method further enhanced classification accuracy; the Average Ensemble achieved the best result at 91.18%. The results show that fusing complementary deep features with ensemble learning yields a stronger, more accurate model than either explicitly using a single feature or using ensemble learning alone.

The confusion matrix was analyzed to further assess the classification performance of the proposed model across all four disease types. A confusion matrix provides insight into correctly and misclassified samples and shows class-wise precision, recall, and F1-score to assess model performance on a per-class basis.

Confusion Matrix of the Proposed Average Ensemble Model

		Predicted Class			
		COVID19	NORMAL	PNEUMONIA	TUBERCULOSIS
Actual Class	COVID19	103	0	1	2
	NORMAL	1	181	50	2
	PNEUMONIA	0	12	378	0
	TUBERCULOSIS	0	0	0	41

Figure 3:Confusion matrix of the proposed Average Ensemble model.

It also illustrates that the proposed model achieved excellent accuracy in recognizing COVID-19 and tuberculosis COVID-19 cases, as indicated by the confusion matrix. The majority of misclassifications were misidentifying the Normal vs. Pneumonia classes, as they can appear very similar radiographically. The class-wise precision, recall, and F1 Scores also indicate that the proposed framework achieved consistent results across all four disease groups.

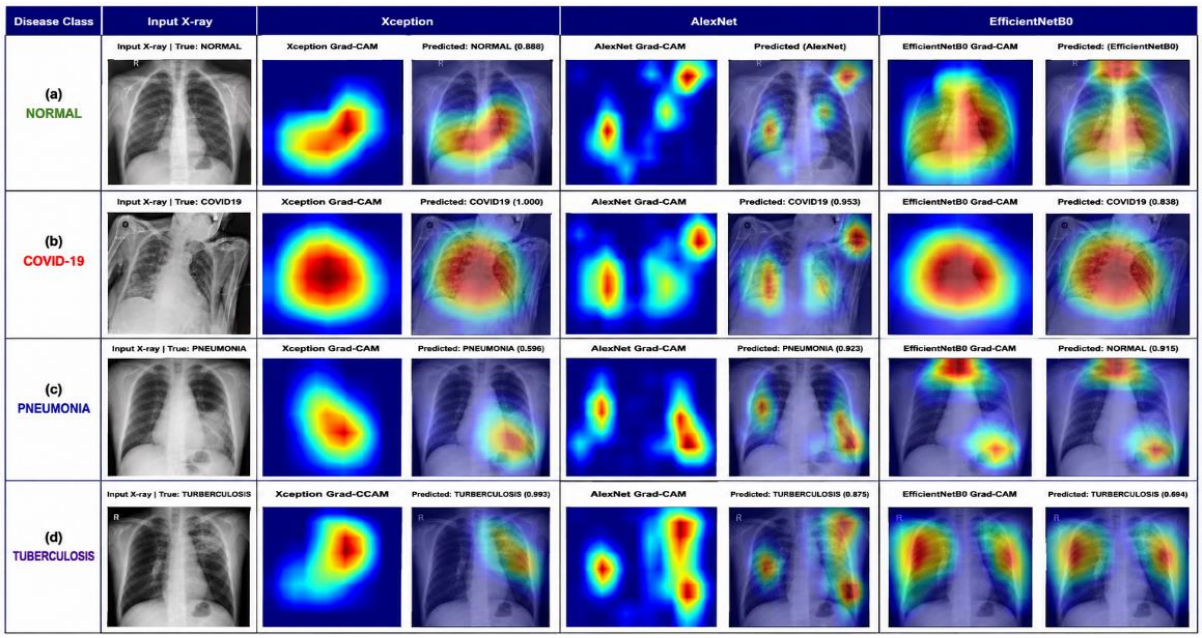

Figure 4: Grad-CAM visualizations for representative chest X-ray images from the four disease classes: (a) Normal, (b) COVID-19, (c) Pneumonia, and (d) Tuberculosis.

The Grad-CAM visualizations for a single chest X-ray image from each of the three disease classes are shown in Figure 4.

The key areas of focus for the CNN models during prediction are highlighted. The models primarily focused on the atypical areas of the lungs for each COVID-19 and Pneumonia case, where infection occurred. The top part of the lungs in the image of Tuberculosis was the focus, highlighting places where the lesions are visible. The Normal chest X-ray had the models allocated to areas of healthy lung tissue rather than any areas of abnormality. These visualizations indicate that the models had acquired useful features from these images and made their predictions on relevant parts of the lungs in a manner suitable for clinical use. As the proposed model is an ensemble of three CNN backbones (Xception, AlexNet, and EfficientNet-B0), Grad-CAM results from each backbone are shown before the ensemble prediction to illustrate the attention regions learned by each backbone.

The Average Ensemble performed well in classification across all four disease classes, as shown in Table 4. If we consider the results, the highest recall was achieved in the Tuberculosis class, with a 98% success rate, meaning that everything marked in this class was correctly identified. Additionally, it achieved good classification results, with a precision of 91.11%, an accuracy of 95.35%, and an F1 score of 95.35%. The model also achieved good accuracy for the COVID-19 and pneumonia classes, reflecting the meaningful features it learned across various chest diseases. Further, the macro-average F1-score of 92.63% indicates that the proposed system performed consistently across all four classes.

TABLE 4: Class-wise precision, recall, and F1-score of the proposed Average Ensemble model.

Model	Class	Precision	Recall	F1-Score
Average Ensemble	Normal	0.937824	0.773504	0.847775
Average Ensemble	COVID-19	0.990385	0.971698	0.980952
Average Ensemble	Pneumonia	0.881119	0.969231	0.923077
Average Ensemble	Tuberculosis	0.91111	1.00	0.953488
Average Ensemble	Macro Average	0.930110	0.928608	0.926323
Average Ensemble	Weighted Average	0.914946	0.911803	0.909797

To compare the proposed Average Ensemble model with the individual CNN models, a statistical significance test was conducted to determine whether the proposed model's performance improvement is statistically meaningful. The same test samples were used, and a significance level of 0.05 was employed in the analysis. The p-values here indicate whether these performance differences are statistically significant.

The results. The statistical significance analysis is presented in Table 5. The average Ensemble model performs significantly better than Gabor Features Only and EfficientNet-B0, with p-values < 0.05. Parts of these differences were not statistically significant between Average Ensemble and the models Xception and AlexNet. These results demonstrate that the proposed model achieves the highest overall accuracy; however, the performance gain over the strongest individual CNN models is also rather modest.

TABLE 5: Statistical significance analysis comparing the proposed Average Ensemble model with the baseline models.

Baseline Model	Proposed Model	P Value	Significant (P<0.05)
Gapour Feature Only	Average Ensemble	$2.0422e^{-29}$	True
Efficient Net-B0	Average Ensemble	$5.9735e^{-05}$	True
AlexNet	Average Ensemble	$5.9664e^{-01}$	False
Xception	Average Ensemble	$2.62437e^{-01}$	False

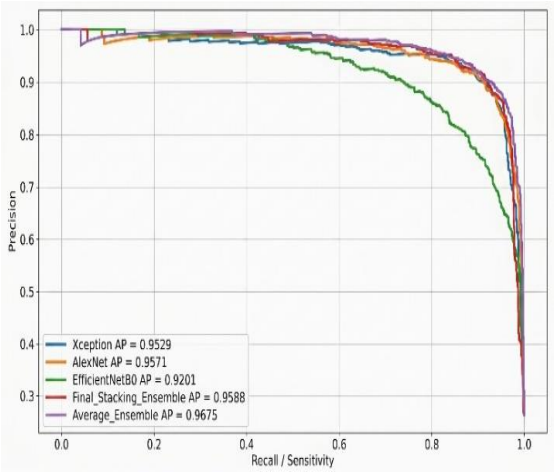

Figure 5: Precision–Recall curves of the evaluated models averaged over six-fold stratified cross-validation.

The precision and recall curve, as shown in Figure 5, also confirmed previous results. Here, Average Ensemble achieved the best AP of 0.9675. The Stacking Logistic Regression Ensemble was the second-highest, with an AP of 0.9588. Blended CNN models achieved the best AP value of 0.9571 for individual CNN models.

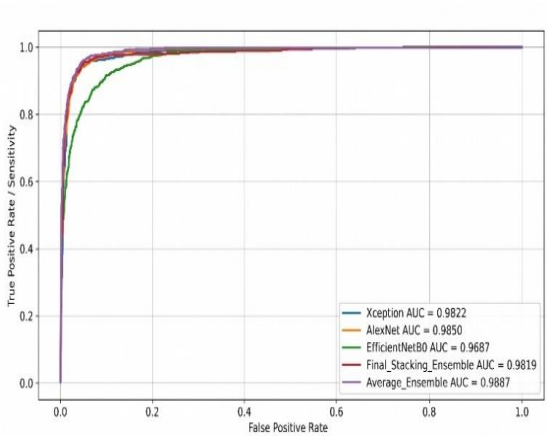
True Positive Rate / Sensitivity
False Positive Rate
Xception AUC = 0.9822
AlexNet AUC = 0.9850
EfficientNetB0 AUC = 0.9687
Final_Stacking_Ensemble AUC = 0.9819
Average_Ensemble AUC = 0.9887

Figure 6: Receiver Operating Characteristic (ROC) curves of the evaluated models.

Micro-average ROC curves of all tested algorithms are given in Fig. 6. The Average Ensemble algorithm achieved an AUC value of 0.9887, while AlexNet achieved an AUC value of 0.9850. The Xception model achieved an AUC of 0.9822, while the Stacking Logistic Regression Ensemble algorithm and EfficientNet-B0 achieved AUCs of 0.9819 and 0.9687, respectively. Therefore, the proposed ensemble model can classify chest X-ray images with satisfactory results in four classes: Normal, COVID-19, Pneumonia, and Tuberculosis by means of the ROC and PR analysis.

Table 6 compares the proposed model with the other related studies presented by various authors. Most of the previously published research was based on binary tuberculosis classification, whereas the present study uses four classes, making the problem statement more complex. In any case, the proposed algorithm could have obtained 91.18% accuracy.

TABLE 6: Comparison of the Proposed Method With Existing Studies.

Ref.	Method	Accuracy (%)
Rajaraman et al. (2018) [1]	Stacked Generalization Ensemble	84
Khatibi et al. (2021) [2]	CNN + Complex Network + Stacked Ensemble	89.36
Ayaz et al. (2021) [4]	Hybrid Feature Descriptor + Ensemble Learning	90.42
Rajaraman & Antani (2020) [8]	Modality-Specific CNN Ensemble	90.13
Proposed Method	-	91.18

DISCUSSION

We compared the performance of both ensembles with that of the individual CNN models, and both ensembles outperformed the CNN models. Still, the Average Ensemble had slightly higher accuracy than the Stacking Logistic Regression Ensemble. This may be because Average Ensemble is a simple, balanced method that equally fuses the prediction probabilities of all base models. This will help minimize prediction variance and stabilize the final prediction output. Instead, the predictions of the Stacking Ensemble are combined using another meta-classifier. The overall performance of the meta-classifier might deteriorate slightly, as the data set is not very large and the classes are far from balanced. Consequently, the Average Ensemble achieved higher consistency in classification results in the four disease classes.

Based on the results, the framework classified chest X-ray images into the Normal, COVID-19, Pneumonia, and Tuberculosis groups. The individual CNN models were compared for accuracy, with AlexNet ranking first, followed by Xception, and EfficientNet-B0 ranking last. Future studies have also reported that learning features using transfer learning with CNN models can be achieved on chest X-ray images[5,7,8].

All three CNN models have distinct learning capacities, which makes them suitable for feature fusion. All these technologies can learn rich spatial information, while AlexNet delivers reliable low-level image features, and EfficientNet-B0 provides efficient feature extraction with fewer model parameters. The framework can extract complementary information that might not be extracted by a single CNN model by combining these models.

The ensemble models also showed further improvement in classification accuracy over the individual CNN models, consistent with the results presented[1,2]. In addition, using Gabor texture features provided a further increase in classification accuracy. The result is consistent with earlier work in which the use of Gabor texture features led to improved results[4,12].

The models' discriminatory power in differentiating Tuberculosis cases was validated using Precision–Recall and ROC curves, as well as the high AP and AUC for the model ensembles. Furthermore, some classification difficulties arose because the X-ray appearances of Normal and Pneumonia were very similar, as noted earlier in the references[7,9].

Compared to previous works that focus only on classifying Tuberculosis detection into two classes [1-6], this work considers a 4-class classification (Normal, COVID-19, Pneumonia, and Tuberculosis). The methods discussed above can be integrated to develop a robust approach to classifying chest diseases.

There are, however, some limitations to the proposed work. It is worth mentioning that the dataset used here is imbalanced with respect to the PNEUMA (Pneumonia) class. Furthermore, the experiments were conducted using a single dataset. An evaluation on multi-center datasets not included in the above work should be conducted to assess the efficiency of the developed method.

One constraint is that the public patient information is not stored with identifiers. However, it was not possible to confirm the independence of the training and validation data sets at the patient level. Further tests of the proposed framework are to be conducted using patient-level datasets to gain deeper insights into its generalization performance.

CONCLUSION

This study aimed to develop a hybrid ensemble learning method to classify chest X-rays into four categories: Normal, COVID-19, Pneumonia, and Tuberculosis. The proposed approach uses deep features from Xception, AlexNet, and EfficientNet, along with Gabor texture features and ensemble learning.

It showed that the ensemble model outperforms individual CNN models in accuracy, AP, and AUC, achieving 91.18% accuracy, 0.9675 AP, and 0.9887 AUC. Results showed that deep learning, texture features, and ensemble learning methods are useful in improving chest disease classification. Larger datasets and more sophisticated explainable AI techniques will be the subject of future studies.

Author Contributions

ARKT: Conceived and designed the study, supervised the research, and contributed to manuscript drafting.

MLURS: Assisted in study design, data collection, and data analysis.

FS: Contributed to the literature review, data interpretation, and manuscript preparation.

MHZ: Assisted in data collection, data entry, and manuscript editing.

SAK: Critically revised the manuscript, provided overall guidance, and approved the final version.

Grant Support and Funding Disclosure
None.

Conflict of Interests
None.

REFERENCES

1. Rajaraman S, Antani S, Candemir S, Xue Z, Alderson P, Kohli M, et al. A novel stacked generalization of models for improved TB detection in chest radiographs. 2018. doi: https://doi.org/10.1201/9780429029417
2. Khatibi T, Shahsavari A, Farahani A. Proposing a novel multi-instance learning model for tuberculosis recognition from chest X-ray images based on CNNs, complex networks, and stacked ensemble. Phys Eng

Sci Med. 2021 Mar 1;44(1):291–311. https://doi.org/10.1007/s13246-021-00980-w
3. Hwa SKT, Hijazi MHA, Bade A, Yaakob R, Jeffree MS. Ensemble deep learning for tuberculosis detection using chest X-ray and Canny edge detected images. IAES International Journal of Artificial Intelligence. 2019 Dec 1;8(4):429–35. doi: https://doi.org/10.11591/ijai.v8.i4.pp429-435
4. Ayaz M, Shaukat F, Raja G. Ensemble learning-based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors. Phys Eng Sci Med. 2021 Mar 1;44(1):183–94. https://doi.org/10.1007/s13246-020-00966-0
5. Wong A, Lee JRH, Rahmat-Khah H, Sabri A, Alaref A, Liu H. TB-Net: A Tailored, Self-Attention Deep Convolutional Neural Network Design for Detection of Tuberculosis Cases From Chest X-Ray Images. Front Artif Intell. 2022 Apr 7;5. doi: https://doi.org/10.3389/frai.2022.827299
6. Munadi K, Muchtar K, Maulina N, Pradhan B. Image Enhancement for Tuberculosis Detection Using Deep Learning. IEEE Access. 2020;8:217897–907. doi: https://doi.org/10.1109/ACCESS.2020.3041867
7. Mirugwe A, Lillian Tamale M, Nyirenda J. Improving Tuberculosis Detection in Chest X-Ray Images Through Transfer Learning and Deep Learning: Comparative Study of Convolutional Neural Network Architectures [Internet]. https://doi.org/10.2196/66029
8. Rajaraman S, Antani SK. Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs. IEEE Access. 2020;8:27318–26. doi: https://doi.org/10.1109/access.2020.2971257
9. Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, et al. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation, and visualization. IEEE Access. 2020;8:191586–601. doi: https://doi.org/10.48550/arXiv.2007.14895
10. Rajaraman S, Kim I, Antani SK. Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles. PeerJ. 2020;2020(3). https://doi.org/10.7717/peerj.8693
11. Sahlol AT, Elaziz MA, Jamal AT, Damaševičius R, Hassan OF. A novel method for the detection of tuberculosis in chest radiographs using artificial ecosystem-based optimization of deep neural network features—symmetry (Basel). 2020 Jul 1;12(7). https://doi.org/10.3390/sym12071146
12. Win KY, Maneerat N, Hamamoto K, Sreng S. Hybrid learning of hand-crafted and deep-activated features using particle swarm optimization and optimized support vector machine for tuberculosis screening. Applied Sciences (Switzerland). 2020 Sep 1;10(17). https://doi.org/10.3390/app10175749
13. Hwa SKT, Bade A, Hijazi MHA, Jeffree MS. Tuberculosis detection using deep learning and contrast-enhanced canny edge-detected X-Ray images. IAES International Journal of Artificial Intelligence. 2020;9(4):713–20. http://doi.org/10.11591/ijai.v9.i4.pp713-720
14. Msonda P, Uymaz SA, Karaağaç SS. Spatial pyramid pooling in deep convolutional networks for automatic tuberculosis diagnosis. Traitement du signal. 2020 Dec 1;37(6):1075–84. https://doi.org/10.18280/ts.370620
15. Pasa F, Golkov V, Pfeiffer F, Cremers D, Pfeiffer D. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization. Sci Rep. 2019 Dec 1;9(1). https://doi.org/10.1038/s41598-019-42557-4
16. Devnath L, Luo S, Summons P, Wang D. Tuberculosis (TB) classification in chest radiographs using deep convolutional neural networks. https://doi.org/10.3389/fmed.2024.1290729
17. Santosh K, Allu S, Rajaraman S, Antani S. Advances in Deep Learning for Tuberculosis Screening using Chest X-rays: The Last 5 Years Review. J Med Syst. 2022 Nov 1;46(11). https://doi.org/10.1007/s10916-022-01870-8
18. Nafisah SI, Muhammad G. An Explainable Graph Neural Network for Harnessing Long-Range Dependencies in Tuberculosis Classifications in Chest X-Ray Images. Neural Comput Appl. 2024 Jan 1;36(1):111–31. https://doi.org/10.3390/diagnostics15243236
19. Venkatachalam C, Shah P. A Dual-Branch Deep Learning Framework Combining Xception and ResNet for Accurate Lung and Colon Cancer Detection. IEEE Access. 2025;13:131483–97. doi: https://ieeexplore.ieee.org/document/11081468/
20. Singh Bisht A, Ajay A, Karthik R. DeepCRC-Net: An Attention-Driven Deep Learning Network for Colorectal Cancer Classification Using Xception and Efficient Lightweight Local Feature Fusion Networks. IEEE Access. 2025;13:49362–74. https://doi:10.1109/ACCESS.2025.3550004
21. Liu L, Xia K. BTIS-Net: Efficient 3D U-Net for Brain Tumor Image Segmentation. IEEE Access. 2024;12:133392–405. doi: https://doi.org/10.1109/ACCESS.2024.3460797
22. Jain M, Shah A. Anomaly Detection Using Convolutional Neural Networks (CNN). International Journal of Advancements in Computational Technology. 2024;2:12–22. doi: https://doi.org/10.56472/25838628/IJACT-V2I3P102
23. Brima Y, Atemkeng M, Djiokap ST, Ebiele J, Tchakounté F. Transfer learning for the detection and diagnosis of types of pneumonia, including pneumonia induced by COVID-19, from chest X-ray images. doi. 2021;11(8):1480. https://doi.org/10.3390/diagnostics11081480
24. Ahmed MS, Rahman A, AlGhamdi F, AlDakheel S, Hakami H, AlJumah A, et al. Joint diagnosis of pneumonia, COVID-19, and tuberculosis from chest X-ray images: A deep learning approach. doi. 2023;13(15):2562. https://doi.org/10.3390/diagnostics13152562
25. Mwendo I, Gikunda P, Maina A. Deep transfer learning for detecting COVID-19, pneumonia, and tuberculosis using CXR images: A review. arXiv. 2023. Doi : https://arxiv.org/abs/2303.16754
26. Mabrouk A, Díaz Redondo RP, Dahou A, Elaziz MA, Kayed M. Pneumonia detection on chest X-ray images using an ensemble of deep convolutional neural networks. arXiv. 2023. Doi. https://arxiv.org/abs/2312.07965
27. Shahzad S, Sajeela, Shah N, Khan SA. Ensemble deep learning for multi-class chest X-ray classification: Robust detection of pneumonia and tuberculosis. Med Res Arch. 2025;13(11). Doi: https://doi.org/10.18103/mra.v13i11.7065